

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 01:36:22 ; Search time 7202 Seconds
(without alignments)
11521.814 Million cell updates/sec

Title: US-10-080-522-2

Perfect score: 2180

Sequence: 1 ATCTCTGTTCTCTTAGCGT.....TTAAATAAGTTTCGTCG 2180

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gse1:*

9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1626.6	74.6	1700	3	CR614987
2	1611.6	73.9	1685	3	CR606594
3	1608.6	73.8	1682	3	CR620624
4	1604.6	73.6	1678	3	CR594118
5	1601.6	73.5	1675	3	CR625551
6	1559	71.5	1650	3	CR594943
7	1445.2	66.3	1633	3	CR595956
8	1402.8	64.3	1461	3	CR591775
9	919.6	42.2	1060	5	CR384442
10	916.4	42.0	1047	5	CR379012
11	889.2	40.8	1057	5	CR399171
12	887.2	40.7	1140	5	CR358342
13	882.6	40.5	1026	5	CR417901
14	869.6	39.9	1025	5	CR384441
15	862.8	39.6	1076	5	CR358343
16	861.2	39.5	1011	5	CR396113
17	847.2	38.9	1013	5	CR380489
18	845	38.8	1053	1	AL513766
19	841.4	38.6	1060	5	CR380490
20	824	37.8	1009	5	CR324662
21	822.6	37.7	972	5	CR384475
22	803.6	36.9	1023	5	CR379011
23	791.8	36.3	840	4	BI820744
24	789.4	36.2	1115	4	BM548056

25: 776.8 35.6 846 4 BI908417 BI908417 603067363
26: 771.6 35.4 1007 5 BM921904 BM921904 AGENCOURT
c 27: 768.4 35.2 1064 5 BX335528 BX335528 BX335528
28: 755.6 34.7 920 4 BI821983 BI821983 603040009
29: 751.8 34.5 1094 4 BM550009 BM550009 AGENCOURT
30: 749.8 34.4 1054 5 BX378128 BX378128 BX378128
31: 748.8 34.3 752 4 BI521087 BI521087 603081680
32: 746.8 34.3 784 4 BI861812 BI861812 603388835
33: 743.4 34.1 850 4 BI763569 BI763569 603050116
34: 740.2 34.0 894 4 BI910123 BI910123 603067972
35: 737.8 33.8 752 5 BX380144 BX380144 BX380144
36: 736.2 33.8 831 5 BX398302 BX398302 BX398302
37: 728.4 33.4 959 5 BU158208 BU158208 AGENCOURT
38: 727.6 33.4 945 5 BX328120 BX328120 BX328120
39: 727.6 33.4 1049 5 BX335529 BX335529 BX335529
40: 724.8 33.2 751 4 BI907545 BI907545 603065903
41: 722.6 33.1 841 4 BI818929 BI818929 603037421
42: 719.6 33.0 824 4 BI909036 BI909036 603070002
c 43: 708 32.5 799 5 BX380143 BX380143 BX380143
44: 707.4 32.4 721 4 BI761265 BI761265 603044014
c 45: 704.8 32.3 876 5 BX434442 BX434442 BX434442

ALIGNMENTS

RESULT 1

CR614987 1700 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CSODI025YG13 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR614987
VERSION CR614987.1 GI:50495794
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1700)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
2 (bases 1 to 1700)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
JOURNAL - Web : www.genoscope.cns.fr
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES
1. 1700 Location/Qualifiers
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI025YG13"
/tissue type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

Query Match 74.6%; Score 1626.6; DB 3; Length 1700;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1688; Conservative 0; Mismatches 9; Indels 6; Gaps 5;
QY 199 TTCTGGGGCTCCGGGGCGCGAGAGTGTGATCCAGAGGCGCGTCCAGAGCGGAC 258
DB 1 TTCTGGGGCTCCGGGGCGCGAGAGTGTGATCCAGAGGCGCGTCCAGAGCGGAC 60

QY	259	COGGGAGTGTTTCAAGAGCCAGTGACAAAGACAGGGGCCCAAGTCCACCACAGCCATGCA	318
DB	61	COGGGAGTGTTTCAAGAGCCAGTGACAAAGACAGGGGCCCAAGTCCACCACAGCCATGCA	120
QY	319	GACCTGCCCTCGCATTCCTTGGCCACAGCTTTCCAGGCCCTTGGAGCCCTCTCTGTTTTT	378
DB	121	GACCTGCCCTCGCATTCCTTGGCCACAGCTTTCCAGGCCCTTGGAGCCCTCTCTGTTTTT	180
QY	379	GGCTGCTCTTGAAGTGCTCAGATGAAGGCTGGGACAGCCCCATCTGCAAGAGGGGT	438
DB	181	GGCTGCTCTTGAAGTGCTCAGATGAAGGCTGGGACAGCCCCATCTGCAAGAGGGGT	240
QY	439	AGTCTCTGTGTCTTGGGGCGAGAACCGTTCATGCTCTGCAACATCTCCAAAGCCCTTCTC	498
DB	241	AGTCTCTGTGTCTTGGGGCGAGAACCGTTCATGCTCTGCAACATCTCCAAAGCCCTTCTC	300
QY	499	CCATGTCAACATCAAGTGTGTCGCCCAACCGGACAGGAGCGGCATCTTCAATGAGGTGGC	558
DB	301	CCATGTCAACATCAAGTGTGTCGCCCAACCGGACAGGAGCGGCATCTTCAATGAGGTGGC	360
QY	559	TCCAGGCTACTTCTCCCGGACCGCTGGCAGCTCCAGGTTCCAGGAGGGGTGGCACAGCT	618
DB	361	TCCAGGCTACTTCTCCCGGACCGCTGGCAGCTCCAGGTTCCAGGAGGGGTGGCACAGCT	420
QY	619	GGTGATCAAAAGGGCCCGGACCTCCATGTGTGGCTGTATCATGTGGCACTCTGTGGGACA	678
DB	421	GGTGATCAAAAGGGCCCGGACCTCCATGTGTGGCTGTATCATGTGGCACTCTGTGGGACA	480
QY	679	CCAGAGAAATACAGACAAGTCAAGTGGAGGTTTCCAGTGCAGAACCCAGTCCGCCCC	738
DB	481	CCAGAGAAATACAGACAAGTCAAGTGGAGGTTTCCAGTGCAGAACCCAGTCCGCCCC	540
QY	739	TGACACTGGGTTCTGGGCTGTGCGAGCGGTGGTCACTGTGTCTTCACTCTTGTGTGCG	798
DB	541	CGACACTGGGTTCTGGGCTGTGCGAGCGGTGGTCACTGTGTCTTCACTCTTGTGTGCG	600
QY	799	TCTGGTCAATGTTCCGCTGGTACAGGTGCCCTGTTCCAGCAACCGCGGAGNAGAGTT	858
DB	601	TCTGGTCAATGTTCCGCTGGTACAGGTGCCCTGTTCCAGCAACCGCGGAGNAGAGTT	660
QY	859	CTTCTCTCTAGAACCCAGATGAAGTTCGACGCCCTCAGAGCGGGAGCCAGCAGGGCCT	918
DB	661	CTTCTCTCTAGAACCCAGATGAAGTTCGACGCCCTCAGAGCGGGAGCCAGCAGGGCCT	720
QY	919	GAGCAGAGCTCCGCTGAATGTGTGGACCCAGAGTCTCCAGGCCACCCAGAGCCGCTGGC	978
DB	721	GAGCAGAGCTCCGCTGAATGTGTGGACCCAGAGTCTCCAGGCCACCCAGAGCCGCTGGC	780
QY	979	ACTGGTGTTCAAACCTCACCACCTTGAGGCCCTGGAGCTGCTGTGCCCCCCCAACCTTGT	1038
DB	781	ACTGGTGTTCAAACCTCACCACCTTGAGGCCCTGGAGCTGCTGTGCCCCCCCAACCTTGT	839
QY	1039	TTCCATATGCGCGACAGCCATAGCCGCTTGAAGGCGAGAGGACACAGAGAGCCAGCC	1098
DB	840	TTCCATATGCGCGACAGCCATAGCCGCTTGAAGGCGAGAGGACACAGAGAGCCAGCC	899
QY	1099	CTGAGTGCCGACCTTGGGTGGGGGCTGGGTCTCTGCTCCACCCCGAGGGGACAGAC	1158
DB	900	CTGAGTGCCGACCTTGGGTGGGGGCTGGGTCTCTGCTCCACCCCGAGGGGACAGAC	959
QY	1159	ACCGGCTTGTCTGGCAGGCTGGGCTCTGTGTGTCACCACTCTCTGGGTGCTGTGACCCCT	1218
DB	960	ACCGGCTTGTCTGGCAGGCTGGGCTCTGTGTGTCACCACTCTCTGGGTGCTGTGACCCCT	1019
QY	1219	TCCCTTCACCCCCAGGTTCTCAAGCTCTGCTTCTTCAGTTTCCAAATAGAACCCAC	1278
DB	1020	TCCCTTCACCCCCAGGTTCTCAAGCTCTGCTTCTTCAGTTTCCAAATAGAACCCAC	1079
QY	1279	TCACTTCGCGACACCCGACTTACAGGACGATGCCCTCTGCTCTGCTCATCAAACC	1338
DB	1080	TCACTTCGCGACACCCGACTTACAGGACGATGCCCTCTGCTCTGCTCATCAAACC	1139

QY	1339	CACAGACCCGGACTCCCTTTCTGCGACCCAGAGCTGGTCCGGCCCAAGTGTGGGGTCCG	1339
Db	1140	CACAGACCCGGACTCCCTTTCTGCGACCCAGAGCTGGTCCGGCCCAAGTGTGGGGTCCG	1199
QY	1399	CTCTCTCCACTCCCGAGGGCTCCGGCCCAAGTGTGGGGGGGCCCTGCGCGAGCCTCAGACA	1458
Db	1200	CTCTCTCCACTCCCGAGGGCTCCGGCCCAAGTGTGGGGGGGCCCTGCGCGAGCCTCAGACA	1259
QY	1459	CACCTGGAGTTCAAGGCG--TGCGGGGGGCTTTGGGACATACCTGTCTCCCTTGGCTATGAGCAG	1516
Db	1260	CACCTCCAGTTCCAGGGCTGTGGGGGGCTTTGGGCGACATACCTGTCTCCCTTGGCTATGAGCAG	1319
QY	1517	GCTTTGGGGGGCCCTTCCGCGCGGACCCCGGGGGCCGAGGTAGGCTGTGGGGGCTTAGAGG	1576
Db	1320	GCTTTGGGGGGCCCTTCCGCGCGGACCCCGGGGGCCGAGGTAGGCTG--GGGGGCTTAGAGG	1378
QY	1577	CTGGGATGGCTTCCTGGCGCCCAACCGCCAGGGGGGCAAGCGCGAGCGCGCTGGGAGGCGGCG	1636
Db	1379	CTGGGATGGCTTCCTGGCGCCCAACCGCCAGGGGGGCG--AGCGAGGCGCGGCTGGGAGGCGGCG	1437
QY	1637	GCGCGGGCTCGGGCTCGGGGGGTGAGGTGGAGCGCTG--CCTCGGGGGCTGCTCGCGCATCCC	1695
Db	1438	GCGCGGGCTCGGGGTGCGGGGGTGTGAGGTGGAGCGCGCTCGGGGGCTGGAGCGCATCCC	1497
QY	1696	TCAGTCCCTCGCGCCACACCGGGGGTTCGCTCCCTCGCTGGCCCAACCGCACCTGCGAGCCTCTT	1755
Db	1498	TCAGTCCCTCGCGCCACACCGGGGGTTCGCTCCCTCGCTGGCCCAACCGCACCTGCGAGCCTCTT	1557
QY	1756	TGGACCCAGATCTGTTTGTCTGCTTTTGTCTTGTCTGCTGCTGCGCGGGGGCCCTTGTATGCTTT	1815
Db	1558	TGGACCCAGATCTGTTTGTCTGCTTTTGTCTTGTCTGCTGCTGCGCGGGGGCCCTTGTATGCTTT	1617
QY	1816	CATCTGTATGGGTGGAAAAATCACCGGGAATCCCCCTTCAGTTCCTTTGAAAAAGTTCCA	1875
Db	1618	CATCTGTATGGGTGGAAAAATCACCGGGAATCCCCCTTCAGTTCCTTTGAAAAAGTTCCA	1677
QY	1876	TGACTTCGAATATCTGAAATGAAG	1898
Db	1678	TGACTTCGAATATCTGAAATGAAG	1700
RESULT 2			
CR606594			
LOCUS		1685 bp mRNA linear HTC 21-JUL-2004	
DEFINITION		full-length cDNA clone CSODI053YJ02 of Placenta Cot 25-normalized of Homo sapiens (human).	
ACCESSION		CR606594	
VERSION		CR606594.1 GI:50487401	
KEYWORDS		HTC; CNI5T_cDNA.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		1 (bases 1 to 1685)	
TITLE		Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
JOURNAL		Full-length cDNA libraries and normalization	
REMARK		Unpublished Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue 2 (bases 1 to 1685) Genoscope. Direct Submission Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr) - Web : www.genoscope.cns.fr) 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prim end enriched, double-strand cDNA was digested with Not I and clone into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. Location/Qualifiers 1. .1685	
FEATURES			
source			

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1053Y02"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 73.9%; Score 1611.6; DB 3; Length 1685;

Best Local Similarity 99.1%; Pred. No. 0;

Matches 1673; Conservative 0; Mismatches 9; Indels 6; Gaps 5;

```
QY 212 GGGGCGGAGAGTGCATCCAGAGAGCGCGTCCAGAGCGGACCGCGGAGTGTTC 271
DB 1 GGGGCGGCGAGAGTGCATCCAGAGAGCGCGTCCAGAGCGGACCGCGGAGTGTTC 60
QY 272 AAGAGCCAGTGAAGAGGACGAGGCGCCAGAGTCCACAGCGCATCGAGACCTGCCCGCTG 331
DB 61 AAGAGCCAGTGAAGAGGACGAGGCGCCAGAGTCCACAGCGCATCGAGACCTGCCCGCTG 120
QY 332 GCATTCCCTGGCCAGCTTCCAGGCGCTTGGGACCCCTCCTGTTTGGCTGCTCCTTG 391
DB 121 GCATTCCCTGGCCAGCTTCCAGGCGCTTGGGACCCCTCCTGTTTGGCTGCTCCTTG 180
QY 392 AGTGTCTAGATGAAGGCTGGGACAGCCCCATCTGCAGAGAGGGGTAGTCTGTGTCT 451
DB 181 AGTGTCTAGATGAAGGCTGGGACAGCCCCATCTGCAGAGAGGGGTAGTCTGTGTCT 240
QY 452 TGGGCGGAGACACCGTATGCTCTGCAACATCTCCAGCGCTTCCCATGTGCAACATC 511
DB 241 TGGGCGGAGAACCGTATGCTCTGCAACATCTCCAGCGCTTCCCATGTGCAACATC 300
QY 512 AAGCTGCGTCCCGCAGGCGAGAGCGCATCTTCAATGAGTGGCTCCAGGCTACTTC 571
DB 301 AAGCTGCGTCCCGCAGGCGAGAGCGCATCTTCAATGAGTGGCTCCAGGCTACTTC 360
QY 572 TCCCGGAGCGGTGGAGCTCAGGTTTCAAGAGCGGTGGACACAGCTGGTATCAAGGC 631
DB 361 TCCCGGAGCGGTGGAGCTCAGGTTTCAAGAGCGGTGGACACAGCTGGTATCAAGGC 420
QY 632 GCCCGGAGCTCCATGCTGGGCTGTACATGTGGACACCTCGTGGGACACGAGAAATAC 691
DB 421 GCCCGGAGCTCCATGCTGGGCTGTACATGTGGACACCTCGTGGGACACGAGAAATAC 480
QY 692 AGACAAGTACAGCTGAGGATTTCAAGTGCAGAACCCAGTCCGCGCCCTGACACTGGGTT 751
DB 481 AGACAAGTACAGCTGAGGATTTCAAGTGCAGAACCCAGTCCGCGCCCTGACACTGGGTT 540
QY 752 TGGCTGTGCCAGCGGTGTCTACTGCTGTTTCATCTCTTGGTCTGCTCATGTTC 811
DB 541 TGGCTGTGCCAGCGGTGTCTACTGCTGTTTCATCTCTTGGTCTGCTCATGTTC 600
QY 812 GCCTGGTACAGTGGCGCTGTTCCAGCAACCGCGGAGAGAGTCTTCTCTCTAGAA 871
DB 601 GCCTGGTACAGTGGCGCTGTTCCAGCAACCGCGGAGAGAGTCTTCTCTCTAGAA 660
QY 872 CCCAGATGAGGTCGAGCCCTCAGAGCGGAGCCAGAGCGGCTCGAGAGAGCCCTCC 931
DB 661 CCCAGATGAGGTCGAGCCCTCAGAGCGGAGCCAGAGCGGCTCGAGAGAGCCCTCC 720
QY 932 GCTGAATGTGGACCCCGAGCTCCAGGCCCCACCCCAAGCGCGCTGGCACTGGTGTCAA 991
DB 721 GCTGAATGTGGACCCCGAGCTCCAGGCCCCACCCCAAGCGCGCTGGCACTGGTGTCAA 780
QY 992 CCCTCACCACTGGAGCCCTGAGTGTCTGCTCCCGCCCAACCTTGTTCATATGCCGC 1051
DB 781 CCCTCACCACTGGAGCCCTGAGTGTCTGCTCCCGCCCAACCTTGTTCATATGCCGC 839
QY 1052 AGACCCATAGCCCTGCAAGCGAGAGGACACAGGAGAGCCAGCCCTGAGTGGCGAC 1111
DB 840 AGACCCATAGCCCTGCAAGCGAGAGGACACAGGAGAGCCAGCCCTGAGTGGCGAC 899
QY 1112 TTGGGTGGGGGCGCTGCTCTCTGCTCCACCGGAGGGGACAGACACCGGCTTGTTC 1171
```

```
DB 900 TTGGGTGGCGGGGCTTGGGTCTCTCGTCCACCCCGAGGCGACAGACACCGGCTTGTTC 959
QY 1172 CGAGGCTGGGCTCTGTGTCAACCACTCTCGGTGCGTGACAGACCTTCCCTCCACCCC 1231
DB 960 CGAGGCTGGGCTCTGTGTCAACCACTCTCGGTGCGTGACAGACCTTCCCTCCACCCC 1019
QY 1232 CGAGGCTTCCAAAGCTCTGTCTTCAAGTTCCTCAAGTTCCTCAAGTTCCTCAAGTTCCT 1291
DB 1020 CGAGGCTTCCAAAGCTCTGTCTTCAAGTTCCTCAAGTTCCTCAAGTTCCTCAAGTTCCT 1079
QY 1292 ACCGAGTACAGAGACGATGCCCTCTCTCTGCTTCCCTCATCAAAACCCACAGACCGGAC 1351
DB 1080 ACCGAGTACAGAGACGATGCCCTCTCTCTGCTTCCCTCATCAAAACCCACAGACCGGAC 1139
QY 1352 TCCCTTCTGTCACCCAGGCTGCTCGGCGCCAGGTGTGGGGTCCGCTCTCTCACTCC 1411
DB 1140 TCCCTTCTGTCACCCAGGCTGCTCGGCGCCAGGTGTGGGGTCCGCTCTCTCACTCC 1199
QY 1412 CAGGCTCCGCGCCCAAGTGAAGGGGCGCTTCCCGAGCGCTCAGACACACCTGAGTTCAG 1471
DB 1200 CAGGCTCCGCGCCCAAGTGAAGGGGCGCTTCCCGAGCGCTCAGACACACCTGAGTTCAG 1259
QY 1472 GGC--TGGGGGGGCTTGGCACATACCTGTCTCTTGGCTATGAGCAGGCTTGGGGGCGC 1529
DB 1260 GGC--TGGGGGGGCTTGGCACATACCTGTCTCTTGGCTATGAGCAGGCTTGGGGGCGC 1319
QY 1530 TTCCGCGCAGCGCCCGGGGCGAGGTAGGTGCTGGGGGCTTAGAGGCTTGGATGGCTCC 1589
DB 1320 TTCCGCGCAGCGCCCGGGGCGAGGTAGGTGCTGGGGGCTTAGAGGCTTGGATGGCTCC 1378
QY 1590 TGGCCCAACCGCCAGGGGCGAAGCGCAGCGCGGCTTGGAGCGCGCGGCTCGGG 1649
DB 1379 TGGCCCAACCGCCAGGGGCGC--AGCGCAGCGCGGCTTGGAGCGCGCGGCTCGGG 1437
QY 1650 CTGGGGGTGAGGTGAGCGCTG--CTCTCGGGGCTGGTCCGCGATCCCTCAGTCCCTCGGC 1708
DB 1438 CTGGGGGTGAGGTGAGCGCGCTCTCGGGGCTTGGAGCGCGCATCCCTCAGTCCCTCGGC 1497
QY 1709 CACCGGGGTGCTGCTCTGCTGCGCCAGCGCTTGGAGCGCGCGGCTTGGAGCGCGAGATCT 1768
DB 1498 CACCGGGGTGCTGCTCTGCTGCGCCAGCGCTTGGAGCGCGCGGCTTGGAGCGCGAGATCT 1557
QY 1769 GTTCATGCTTTGCTCTGCTCAGTCCGCGGGGCGCTTGTGATGCTTCTCATGTATGGGG 1828
DB 1558 GTTCATGCTTTGCTCTGCTCAGTCCGCGGGGCGCTTGTGATGCTTCTCATGTATGGGG 1617
QY 1829 TGGAAAAATACCGGGAATCCCTTCAAGTTCCTTGTGAAAAAGTTCATGACTCGAATATC 1888
DB 1618 TGGAAAAATACCGGGAATCCCTTCAAGTTCCTTGTGAAAAAGTTCATGACTCGAATATC 1677
QY 1889 TGAATGA 1896
DB 1678 TGAATGA 1685
```

RESULT 3
CR620624LOCUS CR620624 1682 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0D010Y120 of HeLa cells Cot 25-normalized
of Homo sapiens (human).

ACCESSION CR620624

VERSION CR620624.1 GI:50501431

KEYWORDS HTC; CNSLT_cDNA.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1682)
AUTHORS Mammalia; Eukarya; Chordata; Craniata; Vertebrata; Euteleostomi;
Li W.B., Gruber C., Jessee J., and Polayes D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :

LOCUS	CR594118	1678 bp	mrna	linear	HTC 21-JUL-2004							
DEFINITION	full-length cDNA clone CSODI036YH21 of Placentia Cot 25-normalized of Homo sapiens (human).											
ACCESSION	CR594118											
VERSION	CR594118.1	GI:50474925										
KEYWORDS	HTC; CNSLT_cDNA.											
SOURCE	Homo sapiens (human)											
ORGANISM	Homo sapiens											
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.											
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.											
TITLE	Full-length cDNA libraries and normalization											
JOURNAL	Unpublished											
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue											
REFERENCE	2 (bases 1 to 1678)											
AUTHORS	Genoscope.											
TITLE	Direct Submission											
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr											
COMMENT	- Web : www.genoscope.cns.fr) 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime ends enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.											
FEATURES	Location/Qualifiers											
source	1..1678											
	/organism="Homo sapiens"											
	/mol_type="mRNA"											
	/db_xref="taxon:9606"											
	/clone="CSODI036YH21"											
	/tissue_type="Placentia Cot 25-normalized"											
	/plasmid="pCMVSPORT_6"											
ORIGIN												
Query Match	73.6%; Score 1604.6; DB 3; Length 1678;											
Best Local Similarity	99.1%; Pred. No. 0;											
Matches 1666; Conservative	0; Mismatches 9; Indels 6; Gaps 5;											
Qy	209	TCGGGGCGGGAGAACTGCATCCACAGAGAGCGCTCCAGGAGCGACCCGGGAGTGT 268										
Db	1	TCGGGGCGGGAGAACTGCATCCACAGAGAGCGCTCCAGGAGCGACCCGGGAGTGT 60										
Qy	269	TTCAAGAGCCAGTCACAGGACCAGGGGGCCAGTCCACAGGCATGCAGACCTGCCCC 328										
Db	61	TTCAAGAGCCAGTCACAGGACCAGGGGGCCAGTCCACAGGCATGCAGACCTGCCCC 120										
Qy	329	CTGGCATTCCTCGGCCACGTTTCCAGGCCCTCTGGGACCTCTCTGTTTTGGCTGCCTCC 388										
Db	121	CTGGCATTCCTCGGCCACGTTTCCAGGCCCTCTGGGACCTCTCTGTTTTGGCTGCCTCC 180										
Qy	389	TTGAGTGCTCAGAAATGAAGGCTGGGACAGCCCCCATCTGCACAGAGGGGGTAGTCTCTGTG 448										
Db	181	TTGAGTGCTCAGAAATGAAGGCTGGGACAGCCCCCATCTGCACAGAGGGGGTAGTCTCTGTG 240										
Qy	449	TCTTGGGGCGAGAACACCGTTCATGTCCTGCAAACATCTCCAAACGGCTTCTCCCATGTCAAC 508										
Db	241	TCTTGGGGCGAGAACACCGTTCATGTCCTGCAAACATCTCCAAACGGCTTCTCCCATGTCAAC 300										
Qy	509	ATCAAGCTGGTGCCTCCACGGGACGAGCGCCATCTTCAATGAGGTGGCTCCAGGCTAC 568										
Db	301	ATCAAGCTGGTGCCTCCACGGGACGAGCGCCATCTTCAATGAGGTGGCTCCAGGCTAC 360										
Qy	569	TTCTCCCGGACCGCTCGCAGCTCCAGGTTTCAGGGAGCGGTGGCACAGCTGGTGATCAAA 628										
Db	361	TTCTCCCGGACCGCTCGCAGCTCCAGGTTTCAGGGAGCGGTGGCACAGCTGGTGATCAAA 420										
Qy	629	GGGCGCCGGGACTCCCATGCTGGGCTGTACATGTGGCACTCTGTGGGACACACAGAGAAAT 688										
Db	421	GGGCGCCGGGACTCCCATGCTGGGCTGTACATGTGGCACTCTGTGGGACACACAGAGAAAT 480										

QY	689	AACAGACAAGTCA	CGCTGGAGGTTT	CAGGTGCAGAA	CCCCAGTCCG	CGCCCTGACACTG	GGG	748
DB	481	AACAGACAAGTCA	CGCTGGAGGTTT	CAGGTGCAGAA	CCCCAGTCCG	CGCCCTGACACTG	GGG	540
QY	749	TTCTGGGCTGTG	CCAGCGGTGCT	CACATGCTCTT	CATCTCTTTG	GTGCGTCTG	GTCATG	808
DB	541	TTCTGGGCTGTG	CCAGCGGTGCT	CACATGCTCTT	CATCTCTTTG	GTGCGTCTG	TCTGGTCATG	600
QY	809	TTGCGCTGGTA	CAGGTGCGGT	TTTCCAGCAA	CGCGGGAGAA	GAAGTTTCTT	CTCCTCCTTA	868
DB	601	TTGCGCTGGTA	CAGGTGCGGT	TTTCCAGCAA	CGCGGGAGAA	GAAGTTTCTT	CTCCTCCTTA	660
QY	869	GAACCCGAGAT	GAAGTTCGAC	CGCTCAGAGC	GGGAGCCAC	GAGGCGCTTG	AGCAGAGCC	928
DB	661	GAACCCGAGAT	GAAGTTCGAC	CGCTCAGAGC	GGGAGCCAC	GAGGCGCTTG	AGCAGAGCC	720
QY	929	TCCGCTGAACT	GTGCGACCC	CAGACTCCG	AGCCACCC	AAAGGCGCTG	GCGACTGTC	988
DB	721	TCCGCTGAACT	GTGCGACCC	CAGACTCCG	AGCCACCC	AAAGGCGCTG	GCGACTGTC	780
QY	989	AAACCCCTCA	CCACACTTGG	AGCCCTGG	AGCTGTCTG	CTCCCCC	CAACCCCTTGT	1048
DB	781	AAACCCCTCA	CCACACTTGG	AGCCCTGG	AGCTGTCTG	CTCCCCC	CAACCCCTTGT	839
QY	1049	CGCAGACCCAT	PAGCCGCTT	CGAAGCAG	AGAGACAC	GAGAGAGCC	AGCCCTGAGT	1108
DB	840	CGCAGACCCAT	PAGCCGCTT	CGAAGCAG	AGAGACAC	GAGAGAGCC	AGCCCTGAGT	899
QY	1109	ACCTTGGGTG	CGGGGCTTGG	GTCTCTG	CTCCACCC	GGGAGGCAC	GACACCGGCT	1168
DB	900	ACCTTGGGTG	CGGGGCTTGG	GTCTCTG	CTCCACCC	GGGAGGCAC	GACACCGGCT	959
QY	1169	TTGCGAGCTG	GGGCTCTGT	GTGTAC	CCACTCTCT	GGGTGCGT	GTGACAGCC	1228
DB	960	TTGCGAGCTG	GGGCTCTGT	GTGTAC	CCACTCTCT	GGGTGCGT	GTGACAGCC	1019
QY	1229	CCCCAGGCTT	TCCAAAGCT	CTGCTTCT	CAGTTTCC	AAAAATGGA	ACCACTCAC	1288
DB	1020	CCCCAGGCTT	TCCAAAGCT	CTGCTTCT	CAGTTTCC	AAAAATGGA	ACCACTCAC	1079
QY	1289	AGCACCCGACT	TACACGAG	CGCATG	CCCCCTC	CTCTGCGCT	CATCAAA	1348
DB	1080	AGCACCCGACT	TACACGAG	CGCATG	CCCCCTC	CTCTGCGCT	CATCAAA	1139
QY	1349	GACTTCCCTT	TCTGCA	CCACCA	GGTGTG	TCGGCC	CCAGGTG	1408
DB	1140	GACTTCCCTT	TCTGCA	CCACCA	GGTGTG	TCGGCC	CCAGGTG	1199
QY	1409	TCCCAGGGCT	CCGCGCC	CAAGTAG	GAGGGG	CCCTTGC	CGGAGCCT	1468
DB	1200	TCCCAGGGCT	CCGCGCC	CAAGTAG	GAGGGG	CCCTTGC	CGGAGCCT	1259
QY	1469	CAGGCG--	TGGGGGGCT	TGGCACA	TACCTGT	CCCTTGG	GCTATG	1526
DB	1360	CAGGCGTGT	TGGGGGGCT	TGGCACA	TACCTGT	CCCTTGG	GCTATG	1319
QY	1527	CCCTTCCGCG	CACGCC	CGGGGCG	GAGGTAG	GGGTCT	GGGGGCT	1586
DB	1320	CCCTTCCGCG	CACGCC	CGGGGCG	GAGGTAG	GGGTCT	GGGGGCT	1378
QY	1587	TCCTGGCC	CCACCG	CAGGGGG	CAAGCG	CAGGCG	GGGCTG	1646
DB	1379	TCCTGGCC	CCACCG	CAGGGGG	CAAGCG	CAGGCG	GGGCTG	1437
QY	1647	GGGCTGGGG	GGTCA	GTTGGA	CGCTG	-	CCTCCG	1705
DB	1438	GGGCTGGGG	GGTCA	GTTGGA	CGCTG	-	CCTCCG	1497
QY	1706	GGCCAC	CCGGGG	GTGCT	CTCCTGT	GTC	CCACCG	1765
DB	1498	GGCCAC	CCGGGG	GTGCT	CTCCTGT	GTC	CCACCG	1557

Qy	1766	TCCTGTTCAATGCTTTTGTCTTCGTCACATGCGCGGGGGCCCTTTGATGCTTCATCTGTATG	1825
Db	1558	TCCTGTTCAATGCTTTTGTCTTCGTCACATGCGCGGGGGCCCTTTGATGCTTCATCTGTATG	1617
Qy	1826	GGGTGGAAAAATCATCCGGGAATCCCCCTTCAGTTCTTTTGAAAAAGTTCCATGACTCGAAT	1885
Db	1618	GGGTGGAAAAATCATCCGGGAATCCCCCTTCAGTTCTTTTGAAAAAGTTCCATGACTCGAAT	1677

Qy	1886 A 1886
	—
Db	1678 A 1678

RESULT_5	CR625551	1675 bp	mRNA	linear	HTC 21-JUL-2004				
LOCUS	CR625551	full-length cDNA clone CSD01029YE07 of Placenta Cot 25-normalized							
DEFINITION	of Homo sapiens (human).								
ACCESSION	CR625551								
VERSION	CR625551.1	GI:50506358							
KEYWORDS	HTC; CNSLT cDNA.								
SOURCE	Homo sapiens (human)								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.								
	1 (bases 1 to 1675);								
	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.								
	Full-length cDNA libraries and normalization								
JOURNAL	Unpublished								
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ invitrogen Corporation 1600								

Query Match	Score 1601.6;	DB 3;	Length 1675;
Best Local Similarity	99.1%;	Pred. No. 0;	
Matches 1663:	Conservative	0: Mismatches	6: Gaps
		9: Indels	5: Gaps

Qy	270	TCAAGAGCCAGTGACAAGGACCAAGGGGGCCCAAGTCCACCAAGCCATGCAGACCTGCCCCC	329
Dβ	61	TCAAGAGCCAGTGACAAGGACCAAGGGGGCCCAAGTCCACCAAGCCATGCAGACCTGCCCCC	120

Qy	330	TGGGATTCCTGGGCACGTTTCCAGGCCCTTGGGACCTCCTGTTTTGGCTGCCTCCT	389
D _b	121	TGGGATTCCTGGGCACGTTTCCAGGCCCTTGGGACCTCCTGTTTTGGCTGCCTCCT	180

Qy	390	TGAGTGCTCAGAAATGAAGCTGGGACAGCCCCCATCTGCA	CAGAGGGGGTAGTCTCTGTGT	449
Db	181	TGAGTGCTCAGAAATGAAGCTGGGACAGCCCCCATCTGCA	CAGAGGGGGTAGTCTCTGTGT	240

Qy	450	CTTGGGCGAGAA	CA	CCGTCATGCTTCGCAACATCTCCAA	CGCCTTCTCCATGTCAACA	509
Db	241	CTTGGGCGAGAA	CA	CCGTCATGCTTCGCAACATCTCCAA	CGCCTTCTCCATGTCAACA	300
Qy	510	TCAAGCTGCGTCC	CA	CGGCGAGAGCGCCATCTTCAAT	GAGTGGCTCCAGGCTACT	569
Db	301	TCAAGCTGCGTCC	CA	CGGCGAGAGCGCCATCTTCAAT	GAGTGGCTCCAGGCTACT	360
Qy	570	TCTCCCGGAGCG	GTGCAGCTCCAGGTT	CAGGAGGCGTGGCACA	CGTGGTGCATCAAAG	629
Db	361	TCTCCCGGAGCG	GTGCAGCTCCAGGTT	CAGGAGGCGTGGCACA	CGTGGTGCATCAAAG	420
Qy	630	CGCGCCGGGACT	CCCATCTGCTGGGCTG	ACATGTGGACCTCTGTGGGAC	ACACAGAGAAATA	689
Db	421	CGCGCCGGGACT	CCCATCTGCTGGGCTG	ACATGTGGACCTCTGTGGGAC	ACACAGAGAAATA	480
Qy	690	ACAGACAAGTCA	CGCTGGAGGTTTCAGGTC	GACAAACCCAGTC	CGGCCCTGACACTGGGT	749
Db	481	ACAGACAAGTCA	CGCTGGAGGTTTCAGGTC	GACAAACCCAGTC	CGGCCCTGACACTGGGT	540
Qy	750	TCTGGCCTGTGC	ACAGCGGTGGTCACTGCTG	CTTCACTCTTGTGGTGC	CTCTGGTCAATGT	809
Db	541	TCTGGCCTGTGC	ACAGCGGTGGTCACTGCTG	CTTCACTCTTGTGGTGC	CTCTGGTCAATGT	600
Qy	810	TCGCTCGTATCA	GAGTGGCGCTGTTTCC	CAGCAAACGCGGGAGAGAGTTCT	CTCTCCTCTAG	869
Db	601	TCGCTCGTATCA	GAGTGGCGCTGTTTCC	CAGCAAACGCGGGAGAGAGTTCT	CTCTCCTCTAG	660
Qy	870	AAACCCAGATGA	AGGTCGACGCCCTCAG	AGCGGAGCCACGAGCGGCTG	AGCAGAGCCT	929
Db	661	AAACCCAGATGA	AGGTCGACGCCCTCAG	AGCGGAGCCACGAGCGGCTG	AGCAGAGCCT	720
Qy	930	CGGCTGAATGTG	AGCCCAAGTCCGAG	CGCCACCCAGGCGCTGGCACTGGTGTCA		989
Db	721	CGGCTGAATGTG	AGCCCAAGTCCGAG	CGCCACCCAGGCGCTGGCACTGGTGTCA		780
Qy	990	AAACCTCAACCA	TTGAGACGCTGAG	GTGCTGTCCCCCAACCCCTGTTTCC	CATATGCC	1049
Db	781	AAACCTCAACCA	TTGAGACGCTGAG	GTGCTGTCCCCCAACCCCTGTTTCC	CATATGCC	839
Qy	1050	GCAGACCCATAG	CGGCTGCAAGG	CAGAGAGGACAGGAGCGCCCTG	AGTGCCGA	1109
Db	840	GCAGACCCATAG	CGGCTGCAAGG	CAGAGAGGACAGGAGCGCCCTG	AGTGCCGA	899
Qy	1110	CTTTGGGTGCGG	GGCCTGGGTCTCTCGT	CCCAACCGGAGGACAGACACCGGCTTGCT		1169
Db	900	CTTTGGGTGCGG	GGCCTGGGTCTCTCGT	CCCAACCGGAGGACAGACACCGGCTTGCT		959
Qy	1170	TGGCAGGCTGGG	CGCTCTGTGTCA	CCCACTCTGGGTGCTGTG	CAGACCCCTTCCGCTCCACC	1229
Db	960	TGGCAGGCTGGG	CGCTCTGTGTCA	CCCACTCTGGGTGCTGTG	CAGACCCCTTCCGCTCCACC	1019
Qy	1230	CCCCAGGTCTT	CAAGTCTGCTTCTCAG	TTTCCAAATGGAAACCACTCACTCCGCA		1289
Db	1020	CCCCAGGTCTT	CAAGTCTGCTTCTCAG	TTTCCAAATGGAAACCACTCACTCCGCA		1079
Qy	1290	GCAACCGACTTA	CCAGGACGATGCCCCCTCT	GTCCCTCATCAAAACCCACAGACCCGG		1349
Db	1080	GCAACCGACTTA	CCAGGACGATGCCCCCTCT	GTCCCTCATCAAAACCCACAGACCCGG		1139
Qy	1350	ACTTCCCTTCTG	CCACCCAGGCTGTG	CGGCCCCAGGTGTGGGGTCCGCTCTCTCCACT		1409
Db	1140	ACTTCCCTTCTG	CCACCCAGGCTGTG	CGGCCCCAGGTGTGGGGTCCGCTCTCTCCACT		1199
Qy	1410	CCAGGCGCTCCG	CGCCCAAGTGAGGGG	GGCCCTTGC	CGGAGCCTCAGACACACTGGAGTTC	1469
Db	1200	CCAGGCGCTCCG	CGCCCAAGTGAGGGG	GGCCCTTGC	CGGAGCCTCAGACACACTGGAGTTC	1259
Qy	1470	AGGGC--	TTGGGGGGCTTGG	CACATACCTGTG	CCCTTGGCTATGACAGGCTTTGGGGC	1527
Db	1260	AGGGCTGTGGG	GGGCGCTTGGCCACATAC	CTGTCTCCCTTGGCTATGACAGGCTTTGGGGC		1319
Qy	1528	CTTTCCGCGCAG	CCGCCCGGGGCGGAG	TAGGGTCTTGGGGGCTTTAGAGG	CTGGGATGGCT	1587

```
Db      1320  CCTTCCGGGAGCCCGGGGCCAGAGTAGGGTC -GGGGGCTTAGAGGCTGGGATGGCT 1378
Qy      1588  CTTGGCCCCCACCAGCGGGGCAAGCGAGGCGGGGCTGGAGGGCGGGCGGGCGGCTCG 1647
Db      1379  CTTGGCCCCCACCAGCGGGGCC -AGGCGAGGCGGGCTGGAGGGCGGGCGGGCGGCTCG 1437
Qy      1648  GGCTGGGGGGTTCAGTGGACGCTG -CTTCGGGGGCTGGTCGGGCATCCTCAGTCCCTCG 1706
Db      1438  GGCTGGGGGGTTCAGTGGACGCGGCCCTCCGGGGCTGGACGGGCATCCTCAGTCCCTCG 1497
Qy      1707  GCACCCGGGGTTCCTCCCTGTCGCCACCGCACCTCGCGAGCTCTTTGGACCCAGAT 1766
Db      1498  GCCACCCGGGGTTCCTCCCTGTCGCCACCGCACCTCGCGAGCTCTTTGGACCCAGAT 1557
Qy      1767  CTGTTTCATGCTTTTCTTCTGTCCTGTCCTGCGGGGCGGCGGCTTTGATGCTTTCATCTGATGG 1826
Db      1558  CTGTTTCATGCTTTTCTTCTGTCCTGTCCTGCGGGGCGGCGGCTTTGATGCTTTCATCTGATGG 1617
Qy      1827  GGTGAAAAAATCACCGGAATCCCTTCATGCTTTGAAAAAGTTCCATGACTGCAA 1884
Db      1618  GGTGAAAAAATCACCGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTGCAA 1675

RESULT 6
LOCUS   CR594943                1650 bp      mRNA      linear      HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DE014YC09 of Placenta of Homo sapiens (human).
ACCESSION CR594943
VERSION   CR594943.1 GI:50475750
KEYWORDS  HTC; CNSLT.cDNA.
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 1650)
AUTHORS   Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE     Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
JOURNAL   Full-length cDNA libraries and normalization
REMARK    Unpublished
          Contact : Feng Liang Email : fliang@lifetech.com URL :
          http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
          Faraday Avenue
          2 (bases 1 to 1650)
REFERENCE Genoscope.
AUTHORS   Direct Submission
TITLE     Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL   BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
          - Web : www.genoscope.cns.fr)
COMMENT   1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
          end enriched, double-strand cDNA was digested with Not I and cloned
          into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
          was normalized. Library was constructed by Life Technologies, a
          division of Invitrogen.
FEATURES  Location/Qualifiers
          source
            1..1650
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DE014YC09"
            /tissue_type="Placenta"
            /plasmid="pCMVSPORT_6"
ORIGIN
Query Match 71.5%; Score 1559; DB 3; Length 1650;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1637; Conservative 0; Mismatches 10; Indels 12; Gaps 6;
Qy      258  CCCGGGAGTGTTCAGAGCCAGTGCACAGGACCAAGGAGCGGGGCCCAAGTCCACAGCCATGC 317
Db      1      CCAGGGAGTGTTCAGAGCCAGTGCACAGGACCAAGGAGCGGGGCCCAAGTCCACAGCCATGC 60
Qy      318  AGACTGCCCTCGGCAATCCCTGGCCAGCTTTTCCAGGCGCTTGGGACCCCTCTCTGTTTTT 377
```

```
Db      61  AGACTGCCCTCGGCAATCCCTGGCCAGCTTTCCAGGCGCTTGGGACCCCTCTCTGTTTTT 120
Qy      378  TGGCTGCCCTCTCTGAGTGCTCAGAAATGAAGGCTGGGACAGGCCCATCTCTGCACAGAGGGGG 437
Db      121  TGGCTGCCCTCTCTGAGTGCTCAGAAATGAAGGCTGGGACAGGCCCATCTCTGCACAGAGGGGG 180
Qy      438  TAGTCTCTGTCTTTGGGGCGAGAACACCGTCATGTCTCTGCAACATCTCCAAACGCTTCT 497
Db      181  TAGTCTCTGTCTTTGGGGCGAGAACACCGTCATGTCTCTGCAACATCTCCAAACGCTTCT 240
Qy      498  CCCATGTCAAATCAAGCTGCGTCCCAAGCGGCGAGGAGCGCCATCTTCATAGAGTGG 557
Db      241  CCCATGTCAAATCAAGCTGCGTCCCAAGCGGCGAGGAGCGCCATCTTCATAGAGTGG 300
Qy      558  CTCCAGGCTACTTCTCCCGGAGCGGCTGGCAGCTCCAGGTTTCAGGAGGCGTGGCACAGC 617
Db      301  CTCCAGGCTACTTCTCCCGGAGCGGCTGGCAGCTCCAGGTTTCAGGAGGCGTGGCACAGC 360
Qy      618  TGGTGATCAAAAGGCGCGCGGACTCCCATGCTGGGCTGTATGTTGGCACTCTCTGTGGGAC 677
Db      361  TGGTGATCAAAAGGCGCGCGGACTCCCATGCTGGGCTGTATGTTGGCACTCTCTGTGGGAC 420
Qy      678  ACCAGAGAAATAACAGACAAGTCAAGCTGGAGGTTTCAGGTCGAGAACCCAGTCCGCCC 737
Db      421  ACCAGAGAAATAACAGACAAGTCAAGCTGGAGGTTT-----CAGAACCCAGTCCGCCC 474
Qy      738  CTGACACTGGGCTTTGGGCTGTGCGAGCGGTGGTCACTGCTGTCTTCATCTCTTGGTGG 797
Db      475  CCGACACTGGGCTTTGGGCTGTGCGAGCGGTGGTCACTGCTGTCTTCATCTCTTGGTGG 534
Qy      798  CTCTGGTCAATGTTCCGCTGGTACAGGTGCCGCTGTTCCAGCAACCGCGGAGAGAAAGT 857
Db      535  CTCTGGTCAATGTTCCGCTGGTACAGGTGCCGCTGTTCCAGCAACCGCGGAGAGAAAGT 594
Qy      858  TCTTCTCTAGAACCCAGATGAAGTTCGAGCGGCTCAGAGCGGAGGCGCCAGAGGCGCC 917
Db      595  TCTTCTCTAGAACCCAGATGAAGTTCGAGCGGCTCAGAGCGGAGGCGCCAGAGGCGCC 654
Qy      918  TGAGCAGAGCTCCGCTGAATGTGTGGACCCCGAGCTCCGAGGCCACCCCAAGGCGCGTGG 977
Db      655  TGAGCAGAGCTCCGCTGAATGTGTGGACCCCGAGCTCCGAGGCCACCCCAAGGCGCGTGG 714
Qy      978  CACTGGTGTCAAACCCCTCAACACTTTGAGCGCTGGAGCTGTCTGTCCTCCCAACCCCTTG 1037
Db      715  CACTGGTGTCAAACCCCTCAACACTTTGAGCGCTGGAGCTGTCTGT -CCCCCAACCCCTTG 773
Qy      1038  TTTCCATATGCGGAGACCCATAGCGCTCGAGGCGAGGACAGGACAGGAGCGGAGCCAGC 1097
Db      774  TTTCCATATGCGGAGACCCATAGCGCTCGAGGCGAGGACAGGACAGGAGCGGAGCCAGC 833
Qy      1098  CTTGAGTCCGACCTTTGGGTCGGGGCTGGGTCTCTCGTCCCAACCGGAGGCGCACAGA 1157
Db      834  CTTGAGTCCGACCTTTGGGTCGGGGCTGGGTCTCTCGTCCCAACCGGAGGCGCACAGA 893
Qy      1158  CACGGCTTGTGTCAGGCTGGGCTCTGTGTCAACCACTCTGGGTGCTGTCAGAGCC 1217
Db      894  CACGGCTTGTGTCAGGCTGGGCTCTGTGTCAACCACTCTGGGTGCTGTCAGAGCC 953
Qy      1218  TTCCCTCTCAACCCCGGAGTCTTCAGTCTGCTTCCTCAGTTTCCAAAATGGAACAC 1277
Db      954  TTCCCTCTCAACCCCGGAGTCTTCAGTCTGCTTCCTCAGTTTTCAAAATGGAACAC 1013
Qy      1278  CTCACCTCCGAGCACCCGACTTACCAGAGCGCATGCCCTCTCTGCTCTCATCAAAC 1337
Db      1014  CTCACCTCCGAGCACCCGACTTACCAGAGCGCATGCCCTCTCTCTGCTCTCATCAAAC 1073
Qy      1338  CCACAGACCCCGACTCTCTTTCTGCAACCCAGGCTGTGCGGCCCAAGGTGTGGGGTCC 1397
Db      1074  CCACAGACCCCGACTCTCTCTTTCTGCAACCCAGGCTGTGCGGCCCAAGGTGTGGGGTCC 1133
Qy      1398  GCTCTCTCACTCCAGGGCTCCGGGCCAAGTGAAGGGGGCCCTTCCCGGAGCCTTCAGC 1457
```

Db 1134 GCTCTCTCCACTCCAGGCGCTCCGGGCCCAAGTAGAGGGGGCCCTGCGGAGCCTCAGAC 1193
QY 1458 ACATGGAGTTAGGCG--TGGGGGGGCTTGGCACAATACCTGTCCCTTGGTATGAGCA 1515
Db 1194 ACATCCAGTTAGGCGCTGTGGGGGGCTTGGCCACAATACCTGTCCCTTGGTATGAGCA 1253
QY 1516 GGTCTTGGGGGGCTTCCCGCGCAGCCCGGGGGCGAGGTAGGCTGTGGGGGCTTAGAG 1575
Db 1254 GGTCTTGGGGGGCTTCCCGCGCAGCCCGGGGGCGAGGTAGGCTGTGGGGGCTTAGAG 1312
QY 1576 GCTGGAGTGGCTCTGTGGCCCCACCGCCAGGGGGCAAGCCAGCGCGGGCTGGAGCGCGC 1635
Db 1313 GCTGGAGTGGCTCTGTGGCCCCACCGCCAGGGGGC-AGCGCAGCGCGGGCTGGAGCGCGC 1371
QY 1636 GCGGGGGCTCGGGCTGGGGGGTCAAGTGGAGCGCTG-CTTCCGGGGCTGGTCCGCGATCC 1694
Db 1372 GCGCGGGCTCGGGCTGGGGGGTCAAGTGGAGCGCGCTG-CTTCCGGGGCTGGTCCGCGATCC 1431
QY 1695 CTGAGTCCCTCGGCCACCGCGGGGTGCTCCCTCGTGGCCACCGCACCTGCGGAGCCTCT 1754
Db 1432 CTGAGTCCCTCGGCCACCGCGGGGTGCTCCCTCGTGGCCACCGCACCTGCGGAGCCTCT 1491
QY 1755 TTGGACCCAGATCTGTTATGCTTTTGTCTTCTGTCATGCGCGGGGGCTTTGATGCT 1814
Db 1492 TTGGACCCAGATCTGTTATGCTTTTGTCTTCTGTCATGCGCGGGGGCTTTGATGCT 1551
QY 1815 TCATCTGATGGGTGGAAATATCACCGGAATCCCTTCACTGCTTTTGAATGTTCC 1874
Db 1552 TCATCTGATGGGTGGAAATATCACCGGAATCCCTTCACTGCTTTTGAATGTTCC 1611
QY 1875 ATGACTCGAATATCTGAATGAAGAAACAAACCGGACTC 1913
Db 1612 ATGACTCGAATATCTGAATGAAGAAACAAACCGGACTC 1650

RESULT 7

CR595956
LOCUS full-length cDNA clone CS01016YA15 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR595956
VERSION CR595956.1 GI:50476763
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1633)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. .1633
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01016YA15"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

REFERENCE

AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match 66.3%; Score 1445.2; DB 3; Length 1633;
Best Local Similarity 93.1%; Pred. No. 0;
Matches 1622; Conservative 0; Mismatches 8; Indels 112; Gaps 6;
QY 199 TTCTTGGGGCTCCGGGGCGCGAGAGCTGCATCCAGAGGAGCGCTCCAGGAGCGGAC 258
Db 1 TTCTTGGGGCTCCGGGGCGCGAGAGCTGCATCCAGAGGAGCGCTCCAGGAGCGGAC 60
QY 259 CGGGAGTGTGTTTCAAGAGCCAGTGAAGAGCCAGGGGCCCAAGTCCCAACGACCATGCA 318
Db 61 CGGGAGTGTGTTTCAAGAGCCAGTGAAGAGCCAGGGGCCCAAGTCCCAACGACCATGCA 120
QY 319 GACCTGCCCCCTTGGCATTTCCCTGGCCACGTTTCCAGAGCGCTTGGGACCTCTCTTTT 378
Db 121 GACCTGCCCCCTTGGCATTTCCCTGGCCACGTTTCCAGAGCGCTTGGGACCTCTCTTTT 180
QY 379 GGTGCTCCTTGTAGTGTCTAGAAATGAAGGTGGGACAGCCCCATCTGACAGAGGGGT 438
Db 181 GGTGCTCCTTGTAGTGTCTAGAAATGAAGGTGGGACAGCCCCATCTGACAGAGGGGT 240
QY 439 AGTCTCTGTGCTTGGGGCGGAGAACACCGTTCATGCTCTGCAACATCTCCAAAGCCTTCTC 498
Db 241 AGTCTCTGTGCTTGGGGCGGAGAACACCGTTCATGCTCTGCAACATCTCCAAAGCCTTCTC 300
QY 499 CCATGTCAACATCAAGCTGCGTGGCCACGCGGAGAGAGCGCCATCTTCAATGAGGTGCG 558
Db 301 CCATGTCAACATCAAGCTGCGTGGCCACGCGGAGAGAGCGCCATCTTCAATGAGGTGCG 360
QY 559 TCCAGGCTACTTCTCCCGGGACGGCTGGCAGCTCCAGGTTTCAAGGAGCGGTGGCACAGCT 618
Db 361 TCCAGGCTACTTCTCCCGGGACGGCTGGCAGCTCCAGGTTTCAAGGAGCGGTGGCACAGCT 420
QY 619 GGTGATCAAGAGCGCGCGGACTCCCATGCTGGGGGTGTACATGTGGGACCTCGTGGGACA 678
Db 421 GGTGATCAAGAGCGCGCGGACTCCCATGCTGGGGGTGTACATGTGGGACCTCGTGGGACA 480
QY 679 CCAGAGAAATACAGACAAGTCAAGCTCGAGGTTTTCAGGTGCAGAAACCCAGTCCGCCCC 738
Db 481 CCAGAGAAATACAGACAAGTCAAGCTCGAGGTTTTCAGGTGCAGAAACCCAGTCCGCCCC 515
QY 739 TGACACTGGGTTCTGGCCCTGTGCCAGCGGTGGTCACTGCTGTCTTCTTCTTCTGTCGC 798
Db 516 ----- 515
QY 799 TCTGTGTATGTTGCGCTGGTACAGGTGCGCTGTTCAGAGAACCCCGGGAGAGAGTT 858
Db 516 -----CAGGTGCGCTGTTCAGAGAACCCCGGGAGAGAGTT 554
QY 859 CTTCTCTCTAGAACCCCGAGATGAAGTCCGAGCCCTCAGAGCGGAGCCAGCAGGGCT 918
Db 555 CTTCTCTCTAGAACCCCGAGATGAAGTCCGAGCCCTCAGAGCGGAGCCAGCAGGGCT 614
QY 919 GAGCAGAGCCTCCGCTGAACTGTGACCCCGCAGACTCCGAGGCCACCCCAAGGCCCTGCG 978
Db 615 GAGCAGAGCCTCCGCTGAACTGTGACCCCGCAGACTCCGAGGCCACCCCAAGGCCCTGCG 674
QY 979 ACTGTGTTTAAACCTCACAATCTGGAGCCCTTGAGCTGTCTCTCCCGCCCAACCTTGT 1038
Db 675 ACTGTGTTTAAACCTCACCCTTGGAGCCCTTGAGCTGTCTCTCTCCCGCCCAACCTTGT 733
QY 1039 TTCCATATGCGCAGACACCATAGCCCTGCAAGGAGAGAGGACACAGGAGAGCGCAGCC 1098
Db 734 TTCCATATGCGCAGACACCATAGCCCTGCAAGGAGAGAGGACACAGGAGAGCGCAGCC 793
QY 1099 CTGAGTCCGACCTTGGGTGGCGGGCTGTGGTCTTCTCTCCCGCCCGGAGGCGACAGAC 1158
Db 794 CTGAGTCCGACCTTGGGTGGCGGGCTGTGGTCTTCTCTCCCGCCCGGAGGCGACAGAC 853
QY 1159 ACCGCTTGTGTCAGAGCTGGGCTCTGTGTACCCACTCTCTGGTGGTGGTGGTGGTGGT 1218
Db 854 ACCGCTTGTGTCAGAGCTGGGCTCTGTGTACCCACTCTCTGGTGGTGGTGGTGGTGGTGGT 913

```
QY 1219 TCCCTTCACCCCGAGGCTTCCAGAGCTCTGCTTCTCAGTTTCCAAATGAAACCAACC 1278
|||
Db 914 TCCCTTCACCCCGAGGCTTCCAGAGCTCTGCTTCTCAGTTTCCAAATGAAACCAACC 973
|||
QY 1279 TCACCTCCGACGACCCGAGTACAGGAGCGATGCCCTCTGCTTCCCTCATCAAC 1338
|||
Db 974 TCACCTCCGACGACCCGAGTACAGGAGCGATGCCCTCTGCTTCCCTCATCAAC 1033
|||
QY 1339 CACAGACCGGAGTCTCCCTTTCGCCACCCAGGCTGTGTCGGGCCCGAGGTGGGTCG 1398
|||
Db 1034 CACAGACCGGAGTCTCCCTTTCGCCACCCAGGCTGTGTCGGGCCCGAGGTGGGTCG 1093
|||
QY 1399 CTCTCTCCATCCGAGGCTCCGCGCCCAAGTGAAGGGGCCCTGCGGAGCCTCAGACA 1458
|||
Db 1094 CTCTCTCCATCCGAGGCTCCGCGCCCAAGTGAAGGGGCCCTGCGGAGCCTCAGACA 1153
|||
QY 1459 CACTGGAGTTCAGGGC--TGGGGGGGCTTGGGCACATCTCTCCCTGGCTATGAGCAG 1516
|||
Db 1154 CACTCGAGTTCAGGGCTGTTGGGGGGCTTGGGCACATCTCTCCCTGGCTATGAGCAG 1213
|||
QY 1517 GCTTTGGGGGCCCTTCCGCGGAGAGCCCGGGGGCGAGGTAGGCTTGGGGCTTAGAGG 1576
|||
Db 1214 GCTTTGGGGGCCCTTCCGCGGAGAGCCCGGGGGCGAGGTAGGCTTGGGGCTTAGAGG 1272
|||
QY 1577 CTGGATGGCTCTTGGGCCCAACCGCAGGGGGCAAGCGCAGGCCCGGGCTGGGAGCGGCG 1636
|||
Db 1273 CTGGATGGCTCTTGGGCCCAACCGCAGGGGGC-AGCGCAGGCCCGGGCTGGGAGCGGCG 1331
|||
QY 1637 GGGGGGCTCGGGCTGGGGGTCAGGTGAGCGCTG-CCTCCGGGGCTGTCCGGATCCC 1695
|||
Db 1332 GCGGGGCTCGGGCTGGGGGTCAGGTGAGCGCGCCCTCCGGGGCTGGACCGCATCCC 1391
|||
QY 1696 TCAGTCCCTCGGCGACCCGGGGGTGCTCCCTCGTGCCCAACCGCAGCTGCCGAGCTCTT 1755
|||
Db 1392 TCAGTCCCTCGGCGACCCGGGGGTGCTCCCTCGTGCCCAACCGCAGCTGCCGAGCTCTT 1451
|||
QY 1756 TGGACCCAGATCTGTATGCTTTTGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTT 1815
|||
Db 1452 TGGACCCAGATCTGTATGCTTTTGTCTGTCTGCTGCTGCTGCTGCTGCTGCTTGTATGCT 1511
|||
QY 1816 CATCTGTATGGGTGGGAAATCAGCGGAATCCCTTCACTTCTTGGAAAGTTCCA 1875
|||
Db 1512 CATCTGTATGGGTGGGAAATCAGCGGAATCCCTTCACTTCTTGGAAAGTTCCA 1571
|||
QY 1876 TGACTCGAATATCTGAATGAAGAAAACAAACGACTCAAACTCCAAAGTAGCTCAA 1935
|||
Db 1572 TGACTCGAATATCTGAATGAAGAAAACAAACGACTCAAACTCCAAAGTAGCTCAA 1631
|||
QY 1936 AT 1937
|||
Db 1632 AT 1633
```

```
RESULT 8
CR591775
LOCUS full-length cDNA clone CS0DK010YL19 of HeLa cells Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR591775
VERSION CR591775.1 GI:50472582
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1461)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
```

```
REFERENCE 2 (bases 1 to 1461)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT: 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five clones
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..1461
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK010YL19"
/tissue_type="HeLa cells Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 64.3%; Score 1402.8; DB 3; Length 1461;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1452; Conservative 0; Mismatches 7; Indels 5; Gaps 4;
QY 199 TTCTTGGGGCTCCGGGGCGCGAGAAAGCTTCCATCCAGAGGAGCGCTCCAGAGCGGAC 258
|||
Db 1 TTCTTGGGGCTCCGGGGCGCGAGAAAGCTTCCATCCAGAGGAGCGCTCCAGAGCGGAC 60
|||
QY 259 CCGGAGAGTGTTCAGAGCCAGTGAAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 318
|||
Db 61 CCGGAGAGTGTTCAGAGCCAGTGAAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
|||
QY 319 GACCTGCCCCCTGGGATTCCTCTGGGCGACGTTTCCAGAGGAGGAGGAGGAGGAGGAGGAG 378
|||
Db 121 GACCTGCCCCCTGGGATTCCTCTGGGCGACGTTTCCAGAGGAGGAGGAGGAGGAGGAGGAG 180
|||
QY 379 GGTCTGCTCTTGTAGTGTCTGAGATGAAGGCTGGGAGAGGAGGAGGAGGAGGAGGAGGAG 438
|||
Db 181 GGTCTGCTCTTGTAGTGTCTGAGATGAAGGCTGGGAGAGGAGGAGGAGGAGGAGGAGGAG 240
|||
QY 439 AGTCTCTGTGTCTGGGGCGGAGAACACCGTCTATGCTGCTGCAACATCTCCAGAGGAGGAG 498
|||
Db 241 AGTCTCTGTGTCTGGGGCGGAGAACACCGTCTATGCTGCTGCAACATCTCCAGAGGAGGAG 300
|||
QY 499 CCATGTCAACATCAAGCTGCGTGCGCCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 558
|||
Db 301 CCATGTCAACATCAAGCTGCGTGCGCCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
|||
QY 559 TCCAGGCTACTTCTCCCGGGAGCGGTGGCAGCTCCAGGTTTCAGGAGGAGGAGGAGGAGGAG 618
|||
Db 361 TCCAGGCTACTTCTCCCGGGAGCGGTGGCAGCTCCAGGTTTCAGGAGGAGGAGGAGGAGGAG 420
|||
QY 619 GGTGATCAAGAGGCGCGGGAGCTCCATGCTGGGGCTGTACATGTGGGAGGAGGAGGAGGAGGAG 678
|||
Db 421 GGTGATCAAGAGGCGCGGGAGCTCCATGCTGGGGCTGTACATGTGGGAGGAGGAGGAGGAGGAG 480
|||
QY 679 CCAGAGAAATCAACAGAGTCAAGCTGAGGTTTCAGGTGCGAGAACCCAGCTCCGCGCCC 738
|||
Db 481 CCAGAGAAATCAACAGAGTCAAGCTGAGGTTTCAGGTGCGAGAACCCAGCTCCGCGCCC 540
|||
QY 739 TGACACTGGGTTCTGGGCTGTGCCAGCGGTGGTCACTGCTGCTTTCATCTCTTGGTCCG 798
|||
Db 541 CGACACTGGGTTCTGGGCTGTGCCAGCGGTGGTCACTGCTGCTTTCATCTCTTGGTCCG 600
|||
QY 799 TCTGTGATGTTGGCTGGTCAAGGTGCGGCTGTTCAGAGCAACCGCGGAGAGAGAGTT 858
|||
Db 601 TCTGTGATGTTGGCTGGTCAAGGTGCGGCTGTTCAGAGCAACCGCGGAGAGAGAGTT 660
|||
QY 859 CTTCTCTAGAACCCAGAGTGAAGTTCGAGCGCTCAGAGCGGAGGAGGAGGAGGAGGAGGAG 918
|||
Db 661 CTTCTCTAGAACCCAGAGTGAAGTTCGAGCGCTCAGAGCGGAGGAGGAGGAGGAGGAGGAG 720
|||
```

```
Qy 919 GAGCAGAGCCTCGCTGAACTGTGGACCCAGACTCCGAGCCACCCAGGCGCGTGGC 978
Db |||
721 GAGCAGAGCCTCGCTGAACTGTGGACCCAGACTCCGAGCCACCCAGGCGCGTGGC 780
Qy 979 ACTGGTGTCAAACCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTTGT 1038
Db |||
781 ACTGGTGTCAAACCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTTGT 839
Qy 1039 TTCCATATGCGCGAGACCCATAGCCGCTGCAAGGAGAGAGACAGAGAGGACGAC 1098
Db |||
840 TTCCATATGCGCGAGACCCATAGCCGCTGCAAGGAGAGAGACAGAGAGGACGAC 899
Qy 1099 CTGAGTCCGACCTTGGGTGGCGGGGCTTGGGTCTCTCGTCCACCGAGGAGCAGAC 1158
Db |||
900 CTGAGTCCGACCTTGGGTGGCGGGGCTTGGGTCTCTCGTCCACCGAGGAGCAGAC 959
Qy 1159 ACCGGCTTGTCTGGAGGCTGGGCTTGTGTCAACCACTCTCTGGGTGGGTGAGACCT 1218
Db |||
960 ACCGGCTTGTCTGGAGGCTGGGCTTGTGTCAACCACTCTCTGGGTGGGTGAGACCT 1019
Qy 1219 TCCCTCCACCCCGGAGCTTCCAAAGCTCTGCTTCTCAGTTTCCAAATGMAACCC 1278
Db |||
1020 TCCCTCCACCCCGGAGCTTCCAAAGCTCTGCTTCTCAGTTTCCAAATGMAACCC 1079
Qy 1279 TCACCTCCGAGACCCGACTTACAGAGAGCATGCCCTCTCTGCTCTCATCAAAAC 1338
Db |||
1080 TCACCTCCGAGACCCGACTTACAGAGAGCATGCCCTCTCTGCTCTCATCAAAAC 1139
Qy 1339 CACAGACCCGAGCTCTCTTCTGCAACCCAGGCTGGTTCGGGCCAGGCTGGGTCG 1398
Db |||
1140 CACAGACCCGAGCTCTCTTCTGCAACCCAGGCTGGTTCGGGCCAGGCTGGGTCG 1199
Qy 1399 CTCTCTCCACTCCAGGGCTCCGCGCCCAAGTGAGGGGCGCCCTGCGGAGCCTCAGACA 1458
Db |||
1200 CTCTCTCCACTCCAGGGCTCCGCGCCCAAGTGAGGGGCGCCCTGCGGAGCCTCAGACA 1259
Qy 1459 CACTGAGTTCAGGGC--TGGGGGGCTTGGGACATACCTGTCTGCTTGGCTATGAGCAG 1516
Db |||
1260 CACTGAGTTCAGGGCTTGGGGGGCTTGGGACATACCTGTCTGCTTGGCTATGAGCAG 1319
Qy 1517 GCTTTGGGGGCTTCCGCGGAGCCCGGGGGCGAGTAGGCTTGGGGCTTAGAGG 1576
Db |||
1320 GCTTTGGGGGCTTCCGCGGAGCCCGGGGGCGAGTAGGCTTGGGGCTTAGAGG 1378
Qy 1577 CTGGATGCTCTGCTGCGCCACCGCCAGGGGGCAAGCGCAGGCGCGGGCTTGGAGCGGG 1636
Db |||
1379 CTGGATGCTCTGCTGCGCCACCGCCAGGGGGCAAGCGCAGGCGCGGGCTTGGAGCGGG 1437
Qy 1637 GGGGGGCTCGGGCTGGGGGGTCA 1660
Db |||
1438 GGGGGGCTCGGGCTGGGGGGTCA 1461
```

```
RESULT 9
BX384442 1060 bp mRNA linear EST 27-APR-2004
LOCUS BX384442 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DK010Y120 5-PRIME, mRNA sequence.
ACCESSION BX384442
VERSION BX384442.2 GI:46624993
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
1 (bases 1 to 1060)
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30455242.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
```

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8574.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?c=CS0DK010BE10QPI&c=8574.f.

FEATURES

Location/Qualifiers

source

1..1060

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DK010Y120"

/cell_type="HELA CELLS COT 25-NORMALIZED"

/cell_line="HELA"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

```
Query Match 42.2%; Score 919.6; DB 5; Length 1060;
Best Local Similarity 96.2%; Pred. No. 7.9e-207;
Matches 960; Conservative 12; Mismatches 21; Indels 5; Gaps 3;

Qy 198 TTTCTGGGGCTCGGGCGCGGAGAGCTGCATCCAGAGAGCGCGTCAGAGCGGA 257
Db |||
1 TTTCCGGGATCTCGGGCGCGGAGAGCTGCATCCAGAGAGCGCGTCAGAGCGGA 60
Qy 258 CCGGGAGTCTTCAAGAGCCAGTGACAAGAGCAGAGGCGCCAAAGTCCACAGCCATGC 317
Db |||
61 CCGGGAGTCTTCAAGAGCCAGTGACAAGAGCAGAGGCGCCAAAGTCCACAGCCATGC 120
Qy 318 AGACCTGCCCCCTGGCATTTCCCTGCCACGTTTCCAGGCGCTTGGGACCTCTCTGTTT 377
Db |||
121 AGACCTGCCCCCTGGCATTTCCCTGCCACGTTTCCAGGCGCTTGGGACCTCTCTGTTT 180
Qy 378 TGGTGTCTCTTCTGAGTGTCTCAGAAATGAAGCTGGGACAGCCCATCTTGCAAGAGGG 437
Db |||
181 TGGTGTCTCTTCTGAGTGTCTCAGAAATGAAGCTGGGACAGCCCATCTTGCAAGAGGG 240
Qy 438 TAGTCTCTGCTCTTGGGGGAGAACACCGTCTCTGCAACATCTTCCAAACCTTCT 497
Db |||
241 TAGTCTCTGCTCTTGGGGGAGAACACCGTCTCTGCAACATCTTCCAAACCTTCT 300
Qy 498 CCCATGTCAACATCAAGCTGCTGCCACCGGGCAGGAGAGCGCCATCTTCAATGAGGTGG 557
Db |||
301 CCCACGTCAACATCAAGCTGCTGCCACCGGGCAGGAGAGCGCCATCTTCAATGAGGTGG 360
Qy 558 CTCAGGCTACTTCTCCGGGACCGGTGGAGCTCCAGGTTTCAGGGAGCGGTGGACAGC 617
Db |||
361 CTCAGGCTACTTCTCCGGGACCGGTGGAGCTCCAGGTTTCAGGGAGCGGTGGACAGC 420
Qy 618 TGGTGTCTCAAAAGGCGCCGGGACCTCCCATGCTGGGCTGTACATGTGCACCTCTGGGAC 677
Db |||
421 TGGTGTCTCAAAAGGCGCCGGGACCTCCCATGCTGGGCTGTACATGTGCACCTCTGGGAC 480
Qy 678 ACCAGAGAAATACAGACAAAGTCAAGCTGGAGGTTTCAGGTGACAGAACCCAGTCCGCC 737
Db |||
481 ACCAGAGAAATACAGACAAAGTCAAGCTGGAGGTTTCAGGTGACAGAACCCAGTCCGCC 540
Qy 738 CTGACATGGGTCTTGGGCTGTGCGAGGGTGGTCACTGTGCTTCTTCACTCTTGGTGG 797
Db |||
541 CCGACACTGGGTCTTGGGCTGTGCGAGGGTGGTCACTGTGCTTCTTCACTCTTGGTGG 600
Qy 798 CTCTGGTCTATGCTTGGCTGTGAGAGTGGCGCTGTTCAGGAGCAACCGCGGAGAGAGT 857
Db |||
601 CTCTGGTCTATGCTTGGCTGTGAGAGTGGCGCTGTTCAGGAGCAACCGCGGAGAGAGT 660
Qy 858 TCTTCTCTCTAGAACCCAGATGAGGTGCGAGCCCTTCAGAGCGGGAGCCAGCAGGGCC 917
```


LOCUS BX399171 1057 bp mRNA linear EST 29-APR-2004
DEFINITION BX399171 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI070YN15 5-PRIME, mRNA sequence.
ACCESSION BX399171
VERSION BX399171.2 GI:46874721
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1057)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced gi:30617803.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8574.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CSODI070CG08QP1&c=8574.f.

FEATURES

Location/Qualifiers
1..1057
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI070YN15"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 40.8%; Score 889.2; DB 5; Length 1057;
Best Local Similarity 97.4%; Pred. No. 1.3e-199;
Matches 956; Conservative 1; Mismatches 19; Indels 6; Gaps 5;
QY 214 GCGCGGAGAGCTGCATCCAGAGGAGCGCGTCCAGGAGCGGACCCGGAGTGTTCAA 273
DB 1 GCGCGGAGAGCTGCATCCAGAGGAGCGCGTCCAGGAGCGGACCCGGAGTGTTCAA 60
QY 274 GAGCCAGTCAAGGACCAAGGCGGCCAAGTCCCAACAGCCATGCAGACCTGCCCTGGC 333
DB 61 GAGCCAGTCAAGGACCAAGGCGGCCAAGTCCCAACAGCCATGCAGACCTGCCCTGGC 120
QY 334 ATTCCCTGCGCCAGTTCCTCCAGGCGCTGGGACCTCTCTGTTTGGTGCCTCTTGA 393
DB 121 ATTCCCTGCGCCAGTTCCTCCAGGCGCTGGGACCTCTCTGTTTGGTGCCTCTTGA 180
QY 394 TGCTCAGATGAAGGCTGGGACAGCCCAATCTGCACAGAGGGGTAGTCTCTGTGCTTG 453
DB 181 TGCTCAGATGAAGGCTGGGACAGCCCAATCTGCACAGAGGGGTAGTCTCTGTGCTTG 240
QY 454 GGGGAGAACACGTCATGTCCTGCAACATCTCCAAAGCGTTCCTCCCAATGCAACATCAA 513
DB 241 GGGGAGAACACGTCATGTCCTGCAACATCTCCAAAGCGTTCCTCCCAATGCAACATCAA 300
QY 514 GTTGCTGCTGCCAGCGGACGAGCGCCATCTTCAATGAGTGGCTCCAGGCTACTTCTC 573
DB 301 GTTGCTGCTGCCAGCGGACGAGCGCCATCTTCAATGAGTGGCTCCAGGCTACTTCTC 360
QY 574 CCGGAGCGGCTGGCAGCTCCAGGTTTCAGGAGCGGTGGCACAGCTGTGTATCAAGGCGC 633
DB 361 CCGGAGCGGCTGGCAGCTCCAGGTTTCAGGAGCGGTGGCACAGCTGTGTATCAAGGCGC 420

QY 634 CCGGAGCTCCCATCTGCTGGCTGTATCATGTGGCACCTCGTGGGACACACAGAGAATAACAG 693
DB 421 CCGGAGCTCCCATCTGCTGGCTGTATCATGTGGCACCTCGTGGGACACACAGAGAATAACAG 480
QY 694 ACAAGTCAAGCTGGAGAGTTTCAGGTGTCAGAACCCAGTCCGCCCTGACACTGGGTCTTG 753
DB 481 ACAAGTCAAGCTGGAGAGTTTCAGGTGTCAGAACCCAGTCCGCCCTGACACTGGGTCTTG 540
QY 754 GCTGTGCGAGCGGTGGTCACTGTGCTTCACTCTTCTTGTGTCGTCTGGTCAATGTCGC 813
DB 541 GCTGTGCGAGCGGTGGTCACTGTGCTTCACTCTTCTTGTGTCGTCTGGTCAATGTCGC 600
QY 814 CTGTTACAGTGGCCCTGTTCCCAACAGCGGAGAGAGTCTTCTCCCTCTAGAACCC 873
DB 601 CTGTTACAGTGGCCCTGTTCCCAACAGCGGAGAGAGTCTTCTCCCTCTAGAACCC 659
QY 874 CCAGATGAAGGTGCGAGCCCTCAGAGCGGAGCCAGCAGGCGCTTGAGCAGACCTCCGC 933
DB 660 CCAGATGAAGGTGCGAGCCCTCAGAGCGGAGCCAGCAGGCGCTTGAGCAGACCTCCGC 719
QY 934 TGAATCTGGAGCCAGACTCGAGCCCAACCCCAAGGCGCTGGACTGGTGTTCAAACC 993
DB 720 TGAATCTGGAGCCAGACTCGAGCCCAACCCCAAGGCGCTGGACTGGTGTTCAAACC 779
QY 994 CTCACACTTGGAGCCCTGAGCTGTCTCCGCCCAACCTTGTTCATATGCGCAG 1053
DB 780 CTCACACTTGGAGCCCTGAGCTGTCTCCGCCCAACCTTGTTCATATGCGCAG 838
QY 1054 ACCATATGCCCTGCAAGGAGGAGGACACAGGAGCCAGCAGCCTTGAGTGCAGACCTT 1113
DB 839 ACCATATGCCCTGCAAGGAGGAGGACACAGGAGCCAGCAGCCTTGAGTGCAGACCTT 897
QY 1114 GGTGGCGGCGCTGGTCTCTCTCCACCGGAGGCGCACAGACCGGTGCTGCTGGC 1173
DB 898 GGTGGCGGCGG-CTGGGTCTCTCTCTCC--ACCCGAGGCGCACAGACCGGTGCTGCTGGA 954
QY 1174 AGGTGGGCGCTCTGTGTCAACC 1195
DB 955 GGTGGGCTGTGTACCACTC 976

RESULT 12

LOCUS BX358342/c
DEFINITION BX358342 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI036YH21 3-PRIME, mRNA sequence.
ACCESSION BX358342
VERSION BX358342.2 GI:46552100
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1140)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 5, 2003 this sequence version replaced gi:30382244.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8574.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CSODI036CDI11NFI&c=8574.f.

FEATURES

Location/Qualifiers
1..1140
/organism="Homo sapiens"

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01036YH21"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 40.7%; Score 887.2; DB 5; Length 1140;
Best Local Similarity 94.3%; Pred. No. 3.9e-199;
Matches 1000; Conservative 15; Mismatches 33; Indels 12; Gaps 9;

QY 813 CTGTGTACAGTGCCTGCTTCCACCAACGCGGAGAGAAAGTTCTTCTCTTAGAAC 872
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1068 SCTGTGACAGGTGCCGCTTCCMAGAACGCGGAGABAR---TTCCTCTCTTAGAAC 1012
QY 873 CCAGATGAGTGTGAGCCCTCAGAGCGGGAGCCAGAGGGCTGTGAGCAGAGCTCC- 931
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1011 CCAGATGAGTGTGAGCCCTCAGAGCGGGAGCCAGAGGGCTGTGAGCAGAGCTCCG 952
QY 932 GCTGAACGTGTGACCCCACTCCGAGCCCA-CCCCAGGCGCTGGCACTGCTTCAA 990
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
951 GTTGAACGTGTGACCCCACTCCGAGCCCA-CCCCAGGCGCTGGCACTGCTTCAA 892
QY 991 ACCCTCACCACCTTGGAGCCCTGGAGCTGTGTCTCCCAACCTTGTCTTCCATATGCG 1050
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
891 ACCCTCACCACCTTGGAGCCCTGGAGCTGTGT-CCCCCAACCTTGTCTTCCATATGCG 833
QY 1051 CAGACCCATAGCGGCTGCAAGGCGAGAGGACACAGGAGACCCAGCCCTGTAGTCCGAC 1110
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
832 CAGACCCATAGCGGCTGCAAGGCGAGAGGACACAGGAGACCCAGCCCTGTAGTCCGAC 773
QY 1111 CTTGGGTGGCGGGCTGGTCTCTCTCCACCGGAGGGGACACAGACCGGCTGTCT 1170
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
772 CTTGGGTGGCGGGCTGGTCTCTCTCTCCACCGGAGGGGACACAGACCGGCTGTCT 713
QY 1171 GCGAGCTGGGGCTCTGTGTACCCACTCTCGGTGGTGGAGACCTTCCCTCCACCC 1230
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
712 GCGAGCTGGGGCTCTGTGTACCCACTCTCGGTGGTGGAGACCTTCCCTCCACCC 653
QY 1231 CCCAGTCTTCAAGCTGTCTTCTCAGTTTCCAAAATGGAAACCACTTCCCTCCGAG 1290
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
652 CCCAGTCTTCAAGCTGTCTTCTCAGTTTCCAAAATGGAAACCACTTCCCTCCGAG 593
QY 1291 CACCCGACTTACAGGACGATGCCCTCCCTCTGCTCCCTCATCAACCCAGACCCGGA 1350
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
592 CACCCGACTTACAGGACGATGCCCTCCCTCTGCTCCCTCATCAACCCAGACCCGGA 533
QY 1351 CTCTCTTCTGCAACCCAGGCTGTTCGGGCGCCAGGTGTGGGTCCGCTCTCTCACTC 1410
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
532 CTCTCTTCTGCAACCCAGGCTGTTCGGGCGCCAGGTGTGGGTCCGCTCTCTCACTC 473
QY 1411 CCAGGGCTCCCGCCCAAGTGAAGGGGGCCCTGCGGAGACCTCAGACACACTGGAGTTCA 1470
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
472 CCA-GGCTCCGGCCCAAGTGAAGGGGGCCCTGCGGAGACCTCAGACACACTGGAGTTCA 414
QY 1471 GGGC--TGGGGGGCTTGGGACATACCTGTCTCTTGGGTATGAGCAGGCTTTGGGGCC 1528
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
413 GGGCTGTGGGGGGCTTGGGACATACCTGTCTCTTGGGTATGAGCAGGCTTTGGGGCC 354
QY 1529 CTTCCGCGCAGCCCGGGGGCGAGGTAGGGTCTTGGGGCTTGGAGGCTGGGATGGCTC 1588
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
353 CTTCCGCGCAGCCCGGGGGCGAGGTAGGGTC-GGGGGCTTGGAGGCTGGGATGGCTC 295
QY 1589 CTTGGCCCAACCGGCGGGGCAAGCGCAGCGCGGGCTGGAGCGCGCGCGGCTCGG 1648
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
294 CTTGGCCCAACCGGCGGGGCGAGGTAGGGTC-AGCGCAGCGCGGGCTGGAGCGCGCGGCTCGG 236
QY 1649 GCTGGGGGGTCAAGTGAAGG-CTGGCTCCGGGGCTGGTGGCGCATCTCTCAGTCCCTCGG 1707
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

FEATURES
Location/Qualifiers
1..1026
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01014YC09"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT.6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 40.5%; Score 882.6; DB 5; Length 1026;
Best Local Similarity 95.0%; Pred. No. 4.7e-198;
Matches 989; Conservative 9; Mismatches 28; Indels 15; Gaps 8;

QY 258 CCCGGAGTGTTCAGAGCCAGTGCACAGGACCAAGGCGCCCAAGTCCACAGCATGC 317
Db 1 CCCGGAGTGTTCAGAGCCAGTGCACAGGACCAAGGCGCCCAAGTCCACAGCATGC 60
QY 318 AGACCTGCCCTTGGCATTTCCTGCGCCACGTTTCCAGGCGCTTGGGACCTCTCTGTTT 377
Db 61 AGACCTGCCCTTGGCATTTCCTGCGCCACGTTTCCAGGCGCTTGGGACCTCTCTGTTT 120
QY 378 TGGTGTCTCTTGTAGTCTCAGATGAGGCTGGGACAGGCCCATCTGTGACAGAGGGG 437
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 121 TGGTGCCTCTTGTAGTCTCAGAAATGAAGGCTGGGACAGCCCATCTGCACAGAGGGG 180
Qy 438 TAGTCTCTGTCTTGGGGCGGAGAACACCGTATGTCCTCTGGAACATCTCCACGCTTCT 497
Db 181 TAGTCTCTGTCTTGGGGCGGAGAACACCGTATGTCCTCTGGAACATCTCCACGCTTCT 240
Qy 498 CCATGTCAAATCAAGCTGCGTCCGACGGGACGAGAGCGCCATCTTCAATGAGGTGG 557
Db 241 CCATGTCAAATCAAGCTGCGTCCGACGGGACGAGAGCGCCATCTTCAATGAGGTGG 300
Qy 558 CTCAGGCTACTTCTCCGGGACGGCTGGCAGCTCCAGGTTTCAGGAGCGTGGCACAGC 617
Db 301 CTCAGGCTACTTCTCCGGGACGGCTGGCAGCTCCAGGTTTCAGGAGCGTGGCACAGC 360
Qy 618 TGGTATCAAGCGCGCCGGGACTCCATGCTGGGCTGTACATGTGGGACCTCTGGGAC 677
Db 361 TGGTATCAAGCGCGCCGGGACTCCATGCTGGGCTGTACATGTGGGACCTCTGGGAC 420
Qy 678 ACCAGAGAAATACAGACAAGTCAAGCTGAGGTTTCAGGTCAGAAACCCAGTCCGCC 737
Db 421 ACCAGAGAAATACAGACAAGTCAAGCTGAGGTTT-----CAGAACCCAGTCCGCC 474
Qy 738 CTGACACTGGGTTCTGGCCCTGTGCCAGCGGTGTCTACTGTCTGTCTTCACTCTTGGTGG 797
Db 475 CCGACACTGGGTTCTGGCCCTGTGCCAGCGGTGTCTACTGTCTGTCTTCACTCTTGGTGG 534
Qy 798 CTCTGTCATGTTCCGCTGTGTACAGTGCCTGTCTTCCAGCAACCGCGGAGAGAAAT 857
Db 535 CTCTGTCATGTTCCGCTGTGTACAGTGCCTGTCTTCCAGCAACCGCGGAGAGAAAT 594
Qy 858 TCTTCTCTCTAGAACCCAGATGAAGTGTGAGGCTGTGAGGCTGTGAGGCGGAGCCAGCGGGC 917
Db 595 TCTTCTCTCTAGAACCCAGATGAAGTGTGAGGCTGTGAGGCGGAGCCAGCGGGC 654
Qy 918 TGAGCAGAGCCTCCGCTGAACCTGTGGACCCAGACTCCGAGCCCAACCCAGCGCGCTGG 977
Db 655 TGAGCAGAGCCTCCGCTGAACCTGTGGACCCAGACTCCGAGCCCAACCCAGCGCGCTGG 714
Qy 978 CACTGCTGTCAAAACCTCACCCTTGGAGCCCTGTGAGGCTGTGAGGCGGAGCCAGCGCTTG 1037
Db 715 CACTGCTGTGTCAAAACCTCACCCTTGGAGCCCTGTGAGGCTGTGAGGCGGAGCCAGCGCTTG 773
Qy 1038 TTTCCATATGCGCGACACCATAGCCCTGTGAAGCGAGAGGACACAGAGAGCGCAGC 1097
Db 774 TTTCCATATGCGCGACACCATAGCCCTGTGAAGCGAGAGGACACAGAGAGCGCAGC 832
Qy 1098 CTGAGTGCAGACCTTGGTGGCGGGCTTGGTCTCTCTGTCCTCCACCCGAGGGCAGAGA 1157
Db 833 CTGAGTGCAGACCTTGGTGGCGGGCTTGGTCTCTCTGTCCTCCACCCGGA-GGCACAGA 891
Qy 1158 CACCGGCTTGTGGCAGGCTGGGCTCTGTGTACCCACTCTGGGTGGTGGTGGTGGTGGTGG 1217
Db 892 CACAGGCTTGTGTGAGGCTGGG-CTYTGTTGTC-CCACTCTGGGTGGTGGTGGTGGTGGTGG 949
Qy 1218 TTTCCCTTCCACCCCGAGTCTTCCAGAGTCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1277
Db 950 T---CCCTTCCACCCCGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1005
Qy 1278 CTCACCTCCGAGCAGCAGCCGAC 1298
Db 1006 YACTTCGACGACCSRCTTAC 1026
```

```
RESULT 14
BX384441/c
LOCUS
DEFINITION BX384441 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DK010Y120 3-PRIME, mRNA sequence.
ACCESSION BX384441
VERSION BX384441.2 GI:46620049
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

```
REFERENCE
AUTHORS Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30453250.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8574.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DK010BE10NP1&c=8574.f.
```

FEATURES

source

```
1..1025
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK010Y120"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

```
Query Match 39.9%; Score 869.6; DB 5; Length 1025;
Best Local Similarity 94.5%; Pred. No. 5.6e-195;
Matches 976; Conservative 18; Mismatches 29; Indels 10; Gaps 9;
Qy 850 GNAGAGGTTCTTCTCTCTAGAACCCAGTGNAGTCCGAGCCCTCAGAGCGGGAGCCCA 909
Db 1025 GARGAATAATTTTSCBTCTARACCCAGATGAAGTTC-CAGGCTCAGAGC-GGAGSCCA 968
Qy 910 GCAGGCGCTGAGCAGAGCTCCGCTGAACCTGTGAGCCCGACACTCCGAGCCACCCCAAG 969
Db 967 GCAGGCGCTGAGCAGAG-CTCGCTGAACCTGTGMAACCCAGACTCGCAGCCACCCCAA 909
Qy 970 GCGCTGCGACTGTGTTCAAACCTCACCCTTGGAGCCCTGGAGCTGTCTGTCCCGCCC 1029
Db 908 GCGCTGCGACTGTGTTCAAACCTCACCCTTGGAGCCCGGAGCTGTGT-GSCCCC 850
Qy 1030 AACCTTGTTCATATGCGCGAGACCCATAGCCGCTTGCAGAGGACAGACACAGCA 1089
Db 849 AACCTTGTTCATATGCGCGAGACCCATAGCCGCTTGCAGAGGACAGACACAGCA 790
Qy 1090 GAGCAGCGCTGAGTGGCGACCTTGGGTGGCGGGCGCTGGGTCTCTGCTCCACCGGAG 1149
Db 789 GAGCAGCGCTGAGTGGCGACCTTGGGTGGCGGGCGCTGGGTCTCTGCTCCACCGGAG 730
Qy 1150 GGCACAGACACCGGCTTGTGGCAGGCTGGGCTCTGTGTCAACCCACTCTGGGTGGCT 1209
Db 729 GGCACAGACACCGGCTTGTGGCAGGCTGGGCTCTGTGTCAACCACTCTCTGGGTGGCT 670
Qy 1210 GCAGAGCCCTTCCCTTCCACCCCGAGTCTTCCAGAGTCTGCTTCTCAGTTTCCAAAT 1269
Db 669 GCAGAGCCCTTCCCTTCCACCCCGAGTCTTCCAGAGTCTGCTTCTCAGTTTCCAAAT 610
Qy 1270 GGAAACACTCTACCTCCGAGCAGCCGACTTACAGAGGAGCATGCCCTTCCCTCTGCCCT 1329
Db 609 GGAAACACTCTACCTCCGAGCAGCCGACTTACAGAGGAGCATGCCCTTCCCTCTGCCCT 550
Qy 1330 CATCAAAACCCAGACACCGGACTCTCTTCTGCGACCCAGGCTGTCTCGGCGCCAGGTG 1389
Db 549 CATCAAAACCCAGACACCGGACTCTCTTCTGCGACCCAGGCTGTCTCGGCGCCAGGTG 490
```

```
QY 1390 TGGGGTCCGCTCTCTCCACTCCAGGGCTCCGCGGCCAAGTGAAGGGGGCCCTGCGCGAG 1449
Db 489 TGGGGTCCGCTCTCTCCACTCCAGGGCTCCGCGGCCAAGTGAAGGGGGCCCTGCGCGAG 431
QY 1450 CTTAGAGACACACTGAGTTCAGGGC--TGGGGGGGCTTGGACATACCTGTCCTTGGC 1507
Db 430 CTTAGAGACACACTGAGTTCAGGGC--TGGGGGGGCTTGGACATACCTGTCCTTGGC 371
QY 1508 TATGAGCAGGCTTTGGGGCCCTTCCGGGAGAGCCCGGGGGCCGAGGTAGGGTCTGGGG 1567
Db 370 TATGAGCAGGCTTTGGGGCCCTTCCGGGAGAGCCCGGGGGCCGAGGTAGGGTCTGGGG 312
QY 1568 GCTTAGAGCTGGGATGGCTCTGGGCCCAAGCCAGGGGGCAAGCGAGGCCGGGCTGG 1627
Db 311 GCTTAGAGCTGGGATGGCTCTGGGCCCAAGCCAGGGGGCAAGCGAGGCCGGGCTGG 253
QY 1628 GAGGGGGGGGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGG 1687
Db 252 GAGGGGGGGGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGG 193
QY 1688 CGATCCCTCAGTCCCTCGGCCCAAGCCGGGGGCTGGCTCCCTGCGGCCACCGCAGCTGCGG 1747
Db 192 CGATCCCTCAGTCCCTCGGCCCAAGCCGGGGGCTGGCTCCCTGCGGCCACCGCAGCTGCGG 133
QY 1748 AGCTCTTTGAGCCAGATCTGTTCAAGCTTTTGTCTTGTCTGTCACCTGCGGGGGCCCTTT 1807
Db 132 AGCTCTTTGAGCCAGATCTGTTCAAGCTTTTGTCTTGTCTGTCACCTGCGGGGGCCCTTT 73
QY 1808 GATGCTTTCANCTGATGGGTGGAAATACCGGGAATCCCTTTCAGTCTTTTGAAA 1867
Db 72 GATGCTTTCANCTGATGGGTGGAAATACCGGGAATCCCTTTCAGTCTTTTGAAA 14
QY 1868 AAGTTCCATGACT 1880
Db 13 AAGTTCCATGACT 1
```

```
RESULT 15
BX358343 1076 bp mRNA linear EST 23-APR-2004
LOCUS BX358343 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1036YH21 5-PRIME, mRNA sequence.
ACCESSION BX358343
VERSION BX358343.2 GI:46553925
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1076)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30384253.
Contact: Genoscope
Genoscope - Centre National de Sequençage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8574.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0D1036CD11Q1&c=8574.f.
Location/Qualifiers
1..1076
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1036YH21"
```

FEATURES

source

```
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

```
Query Match 39.6%; Score 862.8; DB 5; Length 1076;
Best Local Similarity 96.4%; Pred. No. 2.3e-193;
Matches 935; Conservative 1; Mismatches 28; Indels 6; Gaps 5;
QY 209 TCCGGGGCGCGAGAGCTGCAATCCACAGAGAGCGCTCCAGAGACGCGAGCCCGGAGTGT 268
Db 1 TCCGGGATCGAGAGAGCTGCAATCCACAGAGAGCGCTCCAGAGACGCGAGCCCGGAGTGT 60
QY 269 TTCAGAGCCAGTGCACAGAGACCGGGGCGCAAGTCCCAACGAGCCATCGAGACCTGCCCC 328
Db 61 TTCAGAGCCAGTGCACAGAGACCGGGGCGCAAGTCCCAACGAGCCATCGAGACCTGCCCC 120
QY 329 CTGGCATTTCCCTGGCCACGTTTCCAGGCGCTTGGGACCCCTCTGTTTTTGGCTGCCCTC 388
Db 121 CTGGCATTTCCCTGGCCACGTTTCCAGGCGCTTGGGACCCCTCTGTTTTTGGCTGCCCTC 180
QY 389 TTGAGTGCTCAGAAATGAAGGCTGGGACAGCCCCATCTGCAAGAGGGGGTGTCTCTGTG 448
Db 181 TTGAGTGCTCAGAAATGAAGGCTGGGACAG--CCCATCTGCACAGAGGGGGTGTCTCTGTG 239
QY 449 TCTTGGGGGAGAACACCGTCAATCTGCAACATCTCCAGCGCTTCCCTCCATGTCAAC 508
Db 240 TCTTGGGGGAGAACACCGTCAATCTGCAACATCTCCAGCGCTTCTCCATGTCAAC 299
QY 509 ATCAAGCTGCTGCCCGCCAGCGGAGAGCGCCATCTTCAATGAGGTGGCTCCAGGCTAC 568
Db 300 ATCAAGCTGCTGCCCGCCAGCGGAGAGCGCCATCTTCAATGAGGTGGCTCCAGGCTAC 359
QY 569 TTCTCCGGGACGGCTGGCAGCTCAGGTTTCAAGGAGCGGTGGACAGCTGTGTATCAAA 628
Db 360 TTCTCCGGGACGGCTGGCAGCTCAGGTTTCAAGGAGCGGTGGACAGCTGTGTATCAAA 419
QY 629 GGCGCCGGGACTCCCATGCTGGGCTGTACATGTGGACCTCTGGGGACACAGAGAAAT 688
Db 420 GGCGCCGGGACTCCCATGCTGGGCTGTACATGTGGACCTCTGGGGACACAGAGAAAT 479
QY 689 AACAGACAAGTCAAGCTGGAGGTTTCAAGTGCAGAACCCAGTCCGCCCCGACACTGGG 748
Db 480 AACAGACAAGTCAAGCTGGAGGTTTCAAGTGCAGAACCCAGTCCGCCCCGACACTGGG 539
QY 749 TTCTGGGCTGTGGCAGGGGTGGTCACTGCTCTTTCATCTCTTGGTCTGGTCTATG 808
Db 540 TTCTGGGCTGTGGCAGGGGTGGTCACTGCTCTTTCATCTCTTGGTCTGGTCTATG 599
QY 809 TTCCGCTGTACAGGTGCGGCTGTTCAGAGCAAGCGCGGGAGAGAAAGTCTTCTCTCTA 868
Db 600 TTCCGCTGTACAGGTGCGGCTGTTCAGAGCAAGCGCGGGAGAGAAAGTCTTCTCTCTA 658
QY 869 GAAACCCAGATGAAGTTCGAGCGCTCAGAGCGGAGCGCCAGAGGGGCTCTGAGCAGAGCC 928
Db 659 GAAACCCAGATGAAGTTCGAGCGCTCAGAGCGGAGCGCCAGAGGGGCTCTGAGCAGAGCC 718
QY 929 TCCGCTGAATGTGGAGCCCAAGATCTGAGAGCCCAAGAGCGGCTGGCACTGGTGTTC 988
Db 719 TCCGCTGAATGTGGAGCCCAAGATCTGAGAGCCCAAGAGCGGCTGGCACTGGTGTTC 778
QY 989 AAACCTCTCAGCTTGGAGCCCTGAGGTGTGTCTCCCGCCCAAGCTTGTTCATATGC 1048
Db 779 AAACCTCTCAGCTTGGAGCCCTGAGGTGTGTCTCCCGCCCAAGCTTGTTCATATGC 836
QY 1049 CGCAGACCATAGCGCTTGAAGGAGAGGAGACAGAGGAGCGCCAGCTTGTAGTGGCG 1108
Db 837 CGCAGACCATAGCGCTTGAAGGAGAGGAGACAGAGGAGCGCCAGCTTGTAGTGGCG 895
QY 1109 ACCTTGGGTGGCGGGGCTGGGTCTCTCTGTCCTCCACCGGAGGCGACAGACACCGGCTTGC 1168
```

Db 896 ACCTGGGTGGGGG-CTGGGTCTCTGTCACCGGAGGCACAGACACGGCTTGCTTG 954
QY 1169 TTGCAGGCT 1178
Db 955 CAGGTGGGCT 964

Search completed: March 20, 2005, 06:49:48
Job time : 7213 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 09:49:49 ; Search time 85 Seconds
(without alignments)
1128.430 Million cell updates/sec

Title: US-10-080-522-1

Perfect score: 248

Sequence: 1 MQTCPLAFPGHVSQLGTL.....PLGALELLSPQLFPFYAADP 248

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_16Dec04.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	248	100.0	248	2 AAW48811	AAW48811 K12 prote
2	248	100.0	248	2 AAY21846	AAY21846 Human sig
3	248	100.0	248	4 AAB36658	AAB36658 Human K12
4	248	100.0	248	8 ADJ75366	ADJ75366 Marker ge
5	248	100.0	248	8 ADP24664	ADP24664 PRO poly
6	136	54.8	149	7 ADB36336	ADB36336 Human imm
7	134	54.0	162	6 AAO29896	AAO29896 Human org
8	123	49.6	183	6 AAO29897	AAO29897 Human org
9	92	37.1	101	6 AAO29895	AAO29895 Human org
10	55	22.2	55	3 AAB34671	AAB34671 Gene 35 h
11	55	22.2	107	6 AAO29898	AAO29898 Human org
12	48	19.4	55	3 AAB34672	AAB34672 Human sec
13	31	12.5	43	7 ADB36337	ADB36337 Human imm
14	30	12.1	30	3 AAB34673	AAB34673 Gene 35 h
15	12	4.8	27	3 AAB34674	AAB34674 Human sec
16	8	3.2	294	6 ABU20614	ABU20614 Protein e
17	8	3.2	340	8 ABM80522	ABM80522 Tumour-as
18	8	3.2	384	6 ABU48570	ABU48570 Protein e
19	8	3.2	453	7 ABO82888	ABO82888 Pseudomon
20	8	3.2	691	8 ADS21140	ADS21140 Bacterial
21	7	2.8	9	8 ADK03559	ADK03559 Hepatitis
22	7	2.8	10	8 ADK03571	ADK03571 Hepatitis
23	7	2.8	17	5 AAE21241	AAE21241 Human gen
24	7	2.8	36	4 AAE01367	AAE01367 Human gen
25	7	2.8	36	5 ABG64111	ABG64111 Human alb

ALIGNMENTS

RESULT 1

AAW48811 standard; protein; 248 AA.

AC AAW48811;

DT 26-OCT-1998 (first entry)

DE K12 protein.

XX DNA probe; CD7 HSI DNase hypersensitive site; mRNA northern blot;
KW human erythroleukemic; HEL; K562 cell line; clone; breast cancer;
KW ovarian cancer; malignant; tumour.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 239

FT /note= "encoded by CCAA"

PN WO9822502-A1.

XX 28-MAY-1998.

XX 24-NOV-1997; 97WO-US021517.

XX 22-NOV-1996; 96US-00755559.

XX (UYDU-) UNIV DUKE.

XX Kaufman RE, Slentz-Kesler KA;

XX WPI; 1998-312415/27.

XX N-PSDB; AAV32446.

XX New isolated K12 protein gene - which is over expressed in certain
neoplastic cells, used to develop products for tumour detection and
treatment.

XX Claim 3; Fig 1; 44pp; English.

XX This present sequence represents the K12 protein, the gene for which has
been located on chromosome 17q25. To obtain this genes cDNA sequence a
500 bp DNA probe, which can be located just upstream of the CD7 HSI DNase
CC hypersensitive site, was used against a mRNA northern blot. From this a
CC 1.8kb transcript was detected in the human erythroleukemic cell line HEL.
CC The probe was then used to screen a human erythroleukemic cell line K562
CC cDNA library, from which several clones were identified and isolated that

Human signal-peptide containing protein coding sequences used to treat cancer and immune responses.

Claim 1; Page 79-80; 99pp; English.

The invention provides human signal-peptide containing proteins (SIGP) (AA21841-855) and polynucleotides (AA82076-90) encoding the proteins. A host cell containing a vector comprising SIGP DNA can be used to produce the SIGP protein. The SIGP protein can be used, in conjunction with a pharmaceutical carrier to treat or prevent a cancer. An antagonist of the SIGP protein can be used to treat or prevent a cancer or an immune response. The cancers that can be treated or prevented include sarcomas, adenocarcinomas, leukemias, lymphomas, melanomas, teratocarcinomas, myelomas and cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The immune responses that can be treated or prevented include, AIDS, Addison's disease, adult respiratory distress syndrome, allergies, anemia, asthma, atherosclerosis, bronchitis, cholecystitis, Crohn's disease, ulcerative colitis, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, atrophic gastritis, glomerulonephritis, Grave's disease, gout, hyperosinophilia, irritable bowel syndrome, lupus erythematosus, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, rheumatoid arthritis, scleroderma, Sjogren's syndrome, and autoimmune thyroiditis, complications of cancer, infections, and trauma

Sequence 248 AA;

Query Match 100.0%; Score 248; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. NO. 2.5e-242;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCPPLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSNCISNA 60
DB 1 MOTCPPLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSNCISNA 60

QY 61 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVGGVAQLVTKGARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVGGVAQLVTKGARDSHAGLYMHLV 120

QY 121 GHQNNRQVTLVSGAEPQSDPTGFWPVPVAVTVFILLVALVNFAYRCRCQOREK 180
DB 121 GHQNNRQVTLVSGAEPQSDPTGFWPVPVAVTVFILLVALVNFAYRCRCQOREK 180

QY 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWPDPSEPTPRPLALVFKPSPLGALLELLSPQ 240
DB 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWPDPSEPTPRPLALVFKPSPLGALLELLSPQ 240

QY 241 LFPYAADP 248
DB 241 LFPYAADP 248

RESULT 3
AAB36658
ID AAB36658 standard; protein; 248 AA.
XX AC AAB36658;
XX DT 13-MAR-2001 (first entry)
XX DE Human K12 protein sequence SEQ ID NO:4.
XX KW Human; CD7; K12; cognate ligand; cluster of differentiation; cancer; identification; inhibiting T cell proliferation; HIV; infection;
KW activating natural killer cell proliferation; leukaemia; lymphoma;
KW sepsis; graft versus host disease; autoimmune disease; arthritis;
KW multiple sclerosis; rheumatoid arthritis; psoriatic arthritis; lupus;
KW scleroderma; psoriasis; atopic dermatitis; type 1 diabetes mellitus;
KW Hashimoto's thyroiditis; pernicious anaemia; Addison's disease; uveitis;
KW myasthenia gravis; psoriasis; Guillain-Barre Syndrome; Grave's disease;

constituted a 1.9kb cDNA. This cDNA was designated K12 and was found to have a single open reading frame as well as being in the same orientation as CD7. The K12 gene was found to be expressed in both breast and ovarian cancer cells at a much higher level than any other malignant or normal tissue that was examined, thus enabling the K12 to be a useful protein in tumour detection and treatment

Sequence 248 AA;

Query Match 100.0%; Score 248; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. NO. 2.5e-242;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCPPLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSNCISNA 60
DB 1 MOTCPPLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSNCISNA 60

QY 61 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVGGVAQLVTKGARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVGGVAQLVTKGARDSHAGLYMHLV 120

QY 121 GHQNNRQVTLVSGAEPQSDPTGFWPVPVAVTVFILLVALVNFAYRCRCQOREK 180
DB 121 GHQNNRQVTLVSGAEPQSDPTGFWPVPVAVTVFILLVALVNFAYRCRCQOREK 180

QY 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWPDPSEPTPRPLALVFKPSPLGALLELLSPQ 240
DB 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWPDPSEPTPRPLALVFKPSPLGALLELLSPQ 240

QY 241 LFPYAADP 248
DB 241 LFPYAADP 248

RESULT 2
AA21846
ID AA21846 standard; protein; 248 AA.
XX AC AA21846;
XX DT 20-SEP-1999 (first entry)
XX DE Human signal peptide-containing protein (SIGP) (clone ID 1747327).
XX KW Signal-peptide containing protein; SIGP; human; cancer; immune response;
KW adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma; AIDS;
KW Addison's disease; adult respiratory distress syndrome; allergy; anemia;
KW asthma; atherosclerosis; bronchitis; cholecystitis; Crohn's disease;
KW ulcerative colitis; atopic dermatitis; dermatomyositis; emphysema;
KW diabetes mellitus; atrophic gastritis; glomerulonephritis; gout; trauma;
KW Grave's disease; hyperosinophilia; irritable bowel syndrome; infection;
KW lupus erythematosus; multiple sclerosis; myasthenia gravis; inflammation;
KW osteoarthritis; osteoporosis; pancreatitis; polymyositis; scleroderma;
KW rheumatoid arthritis; Sjogren's syndrome; autoimmune thyroiditis.
XX OS Homo sapiens.
XX KW WO9933981-A2.
XX PD 08-JUL-1999.
XX PF 22-DEC-1998; 98WO-US027598.
XX PR 31-DEC-1997; 97US-00002485.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Lal P, Hillman JL, Corley NC, Guegler KJ, Baughn MR, Sather SK;
XX PI Shah P;
XX DR WPI; 1999-430242/36.
XX DR N-PSDB; AA82081.
XX

KW systemic lupus erythematosus; dermatomyositis; asthma; eczema;
KW atopic dermatitis; contact dermatitis; eczematous dermatitis;
KW seborrheic dermatitis; rhinitis.

OS Homo sapiens.

XX WO200073333-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US014612.

XX 28-MAY-1999; 99US-0136450P.

XX (IMV) IMMUNEX CORP.

XX Lyman SD, Fanslow WC;

XX WPI; 2001-061511/07.

DR N-PSDB; AAC88152.

XX Stimulating intracellular signaling of CD7 comprises contacting a cell
PT expressing CD7 with recombinant K12 protein, the cognate ligand of CD7,
PT to inhibit T cell proliferation and/or activate natural killer cell
PT proliferation.

XX Claim 2; Page 38-39; 42pp; English.

XX The present invention describes a method for stimulating (S) the
CC intracellular signalling of CD (cluster of differentiation) 7 comprising
CC contacting a cell that expresses CD7 with a recombinant K12 protein (I),
CC the cognate ligand of CD7. (S) is useful for inhibiting T cell
CC proliferation and/or activating NK (natural killer) cell proliferation
CC and/or inducing NK toxicity in a mammal which involves administration of
CC K12 protein. It is also used for treating HIV-1 infection, cancer (T cell
CC leukaemia, acute lymphocytic leukaemia, cutaneous T cell lymphoma),
CC bacterial and viral infections, mediated by CD7. In the case of treating
CC T cell leukaemia the soluble K12 protein is covalently attached to a
CC toxin. A disease mediated by CD7 such as sepsis, graft versus host
CC disease due to transplantation, autoimmune diseases, multiple sclerosis,
CC arthritis, rheumatoid arthritis, psoriatic arthritis, scleroderma, lupus,
CC psoriasis, atopic dermatitis, type 1 diabetes mellitus, Hashimoto's
CC thyroiditis, pernicious anaemia, Addison's disease, myasthenia gravis,
CC uveitis, psoriasis, Guillain-Barre Syndrome, Grave's disease, systemic
CC lupus erythematosus and dermatomyositis, asthma, eczema, atopic
CC dermatitis, contact dermatitis, other eczematous dermatides, seborrheic
CC dermatitis, and rhinitis is also treated by administering a K12
CC antagonist (neutralising antibody). The present sequence represents the
CC human K12 protein, which is given in the exemplification of the present
CC invention

XX Sequence 248 AA;

Query Match 100.0%; Score 248; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 2.5e-242;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTCPLAPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVSVSGENTVMSNCISNA 60
DB 1 MTCPLAPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVSVSGENTVMSNCISNA 60
QY 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGMQLQVGGVQALVIKGRDASHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGMQLQVGGVQALVIKGRDASHAGLYMHLV 120
QY 121 GHQRNRQVTLVSGAEPQSDPTGFWPVAVTAVFILLVALVNFAYWRCRCQORREK 180
DB 121 GHQRNRQVTLVSGAEPQSDPTGFWPVAVTAVFILLVALVNFAYWRCRCQORREK 180
QY 181 KFFLLEPQMKVAALRAGAQGGLSASAEIWTDPSEPTPLALVFKPSPGLGALELLSPQ 240
DB 181 KFFLLEPQMKVAALRAGAQGGLSASAEIWTDPSEPTPLALVFKPSPGLGALELLSPQ 240

QY 241 LFPYAADP 248
DB |||||
241 LFPYAADP 248

RESULT 4

ADJ75366

ID ADJ75366 standard; protein; 248 AA.

XX AC ADJ75366;

XX 20-MAY-2004 (first entry)

XX Marker gene related amino acid sequence SEQ ID NO:618.

XX bronchial asthma; chronic obstructive pulmonary disease;

XX respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;

XX gene therapy; marker.

XX Homo sapiens.

XX EPI394274-A2.

XX 03-MAR-2004.

XX 04-AUG-2003; 2003EP-00254857.

XX 06-AUG-2002; 2002JP-00229312.

XX 20-MAR-2003; 2003JP-00077212.

XX (GENO-) GENOX RES INC.

XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;

XX WPI; 2004-193155/19.

XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.

XX Example 11; SEQ ID NO 618; 241pp; English.

XX The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.

XX

SQ Sequence 248 AA;
 Query Match 100.0%; Score 248; DB 8; Length 248;
 Best Local Similarity 100.0%; Pred. No. 2.5e-242; Mismatches 0; Indels 0; Gaps 0;
 Matches 248; Conservative 0;

QY 1 MQTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSNCISNA 60
 DB 1 MQTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSNCISNA 60
 QY 61 FSHVNIKLRAHQESAI FNEVAPGYSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
 QY 121 GHORNRQVTLVSGAEPOSAPDTGFWPVAVTAVFILLVAFVWFAWYRCRCQORREK 180
 DB 121 GHORNRQVTLVSGAEPOSAPDTGFWPVAVTAVFILLVAFVWFAWYRCRCQORREK 180
 QY 181 KFFLLEPQMKVAALRAGAQOGLSRASAEI WTPDSEPTPRPLALVPKPSPLGALLELLSPQ 240
 DB 181 KFFLLEPQMKVAALRAGAQOGLSRASAEI WTPDSEPTPRPLALVPKPSPLGALLELLSPQ 240
 QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 5
 ADP24664
 ID ADP24664 standard; protein; 248 AA.
 AC ADP24664;
 DT 18-NOV-2004 (first entry)
 XX PRO polypeptide SEQ ID NO:1842.
 XX PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
 KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
 KW antiaesthetic; hepatotropic; respiratory; gene therapy; immune system.
 XX Unidentified.
 OS
 XX W02004041170-A2.
 XX
 XX 21-MAY-2004.
 XX 30-OCT-2003; 2003WO-US034312.
 XX 01-NOV-2002; 2002US-0423394P.
 XX (GETH) GENENTECH INC.
 XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
 PI Wu TD;
 XX WPI; 2004-419628/39.
 DR N-PSDB; ADP24663.
 XX New PRO polypeptides and polynucleotides, useful for treating e.g.
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
 PT renal disease, or demyelinating diseases of the central or peripheral
 PT nervous system.
 XX Claim 7; SEQ ID NO 1842; 2940pp; English.
 PS The invention relates to a novel isolated nucleic acid and the PRO
 CC polypeptide encoded by it. A protein of the invention has
 CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
 CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
 CC antiaesthetic, hepatotropic, and respiratory activity. A polynucleotide
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its
 CC agonist, antagonist, or antibody that specifically binds to the

CC polypeptide is useful for treating an immune related disorder such as
 CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
 CC idiopathic inflammatory myopathy, Sjogren's syndrome, autoimmune
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
 CC disease, a demyelinating disease of the central or peripheral nervous
 CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
 CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
 CC disease, infectious or autoimmune chronic active hepatitis, primary
 CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
 CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
 CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
 CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
 CC disease, asthma, allergic rhinitis, atopic dermatitis, food
 CC hypersensitivity, urticaria, an immunologic disease of the lung,
 CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
 CC pneumonitis, a transplantation associated disease, graft rejection or
 CC graft-versus-host disease. The present sequence represents a PRO protein
 CC of the invention.
 XX
 SQ Sequence 248 AA;
 Query Match 100.0%; Score 248; DB 8; Length 248;
 Best Local Similarity 100.0%; Pred. No. 2.5e-242; Mismatches 0; Indels 0; Gaps 0;
 Matches 248; Conservative 0;

QY 1 MQTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSNCISNA 60
 DB 1 MQTCLAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSNCISNA 60
 QY 61 FSHVNIKLRAHQESAI FNEVAPGYSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGYSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
 QY 121 GHORNRQVTLVSGAEPOSAPDTGFWPVAVTAVFILLVAFVWFAWYRCRCQORREK 180
 DB 121 GHORNRQVTLVSGAEPOSAPDTGFWPVAVTAVFILLVAFVWFAWYRCRCQORREK 180
 QY 181 KFFLLEPQMKVAALRAGAQOGLSRASAEI WTPDSEPTPRPLALVPKPSPLGALLELLSPQ 240
 DB 181 KFFLLEPQMKVAALRAGAQOGLSRASAEI WTPDSEPTPRPLALVPKPSPLGALLELLSPQ 240
 QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 6
 ADB36336
 ID ADB36336 standard; protein; 149 AA.
 AC ADB36336;
 XX 04-DEC-2003 (first entry)
 XX Human immune response associated protein IRAP-7 SEQ ID NO:7.
 XX human; immune response associated protein; IRAP; anti-HIV; antiallergic;
 KW antianemic; antiaesthetic; antiarteriosclerotic; dermatological;
 KW antiinflammatory; antidiabetic; nephrotropic; antithyroid; thyromimetic;
 KW immunosuppressive; antirheumatic; antiarthritic; osteopathic;
 KW antibacterial; virucide; antiparasitic; protozoacide; fungicide;
 KW cerebroprotective; neuroprotective; nootropic; antiparkinsonian;
 KW antipsoriatic; cytostatic; cardiant; gene therapy;
 KW immune system disorder; neurological disorder; developmental disorder;
 KW muscle disorder; cell proliferative disorder; AIDS; allergy; anaemia;
 KW asthma; atherosclerosis; contact dermatitis; Crohn's disease; diabetes;
 KW glomerulonephritis; Grave's disease; Hashimoto's thyroiditis;
 KW multiple sclerosis; rheumatoid arthritis; osteoporosis;
 KW systemic lupus erythematosus; infection; stroke; Alzheimer's disease;
 KW Parkinson's disease; psoriasis; cancer; cardiomyopathy.

OS Homo sapiens.
 XX WO2003074726-A2.
 PN 12-SEP-2003.
 PD 28-FEB-2003; 2003WO-US006307.
 XX 01-MAR-2002; 2002US-0361088P.
 PR 27-MAR-2002; 2002US-0368494P.
 PR 10-MAY-2002; 2002US-0379876P.
 PR 28-JUN-2002; 2002US-0392641P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Chawla NK, Becha SD, Jin P, Tran UK, Kable AE, Lee S;
 PI Marquis JP, Emerling BM, Ison CH, Ramkumar J, Chang H;
 XX WPI; 2003-722079/68.
 DR N-PSDB; ADB36352.
 XX New human immune response associated proteins and polynucleotides, useful
 PT for diagnosing, preventing or treating diseases or conditions associated
 PT with aberrant enzyme expression, e.g. cancer, AIDS, atherosclerosis or
 PT stroke.
 XX Claim 1; Page 143; 158pp; English.
 PS ADB36346 to ADB36361 encode the human immune response associated proteins
 CC given in ADB36330 to ADB36345, designated IRAP-1 to IRAP-16. IRAP
 CC sequences have anti-HIV, anti-allergic, antianemic, antiasthmatic,
 CC antiarteriosclerotic, dermatological, anti-inflammatory, antidiabetic,
 CC nephrotropic, antithyroid, thyromimetic, immunosuppressive,
 CC antirheumatic, antiarthritic, osteopathic, antibacterial, virucide,
 CC antiparasitic, protozoacide, fungicide, cerebroprotective,
 CC neuroprotective, neurotropic, antiparkinsonian, antipsoriatic, cytostatic
 CC and cardiant activities, and can be used in gene therapy. The IRAP
 CC proteins and polynucleotides can be used in diagnosing, preventing or
 CC treating diseases or conditions associated with the decreased expression
 CC or overexpression of IRAP, such as immune system, neurological,
 CC developmental, muscle or cell proliferative disorders. The disorders may
 CC include AIDS, allergies, anaemia, asthma, atherosclerosis, contact
 CC dermatitis, Crohn's disease, diabetes, glomerulonephritis, Grave's
 CC disease, Hashimoto's thyroiditis, multiple sclerosis, rheumatoid
 CC arthritis, osteoporosis, systemic lupus erythematosus, infections (e.g.
 CC bacterial, viral, fungal, parasitic or protozoal), stroke, Alzheimer's
 CC disease, Parkinson's disease, psoriasis, cancer and cardiomyopathy. They
 CC are also useful in assessing the effects of exogenous compounds on the
 CC expression of nucleic acid and amino acid sequences of IRAP. The IRAP or
 CC its fragments are useful in screening compounds for effectiveness as
 CC agonist or antagonist of the polypeptides, or in altering the expression
 CC of the target polynucleotide and compounds that specifically bind to or
 CC modulate the activity of the polypeptide.
 XX Sequence 149 AA;
 SQ
 Query Match 54.8%; Score 136; DB 7; Length 149;
 Best Local Similarity 100.0%; Pred. No. 4.2e-129;
 Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNINA 60
 DB 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNINA 60
 QY 61 FSHVNIKILRAHGOEAIENEVAPGYFSDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKILRAHGOEAIENEVAPGYFSDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
 QY 121 GQRNNRQVTLVSGA 136
 DB 121 GQRNNRQVTLVSGA 136

RESULT 7
 AAO29896
 ID AAO29896 standard; protein; 162 AA.
 XX AC AAO29896;
 XX DT 03-SEP-2003 (first entry)
 XX Human organelle-associated protein (ORGA)-9.
 DE Human; organelle-associated protein; ORGA; cell proliferative disorder;
 KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
 KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
 KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
 KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
 KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
 KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
 KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
 KW infection; urethritis; prostatitis.
 XX Homo sapiens.
 OS WO2003044171-A2.
 PN 30-MAY-2003.
 PD 15-NOV-2002; 2002WO-US036807.
 PF 16-NOV-2001; 2001US-0332384P.
 PR 13-DEC-2001; 2001US-0341187P.
 PR 23-JAN-2002; 2002US-0351151P.
 PR 27-FEB-2002; 2002US-0360269P.
 PR 05-APR-2002; 2002US-0370637P.
 PR 14-JUN-2002; 2002US-0388946P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
 PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
 PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
 PI Hafalia AJA, Bulloch S;
 XX WPI; 2003-457603/43.
 DR N-PSDB; AAL60549.
 XX New organelle-associated proteins and polynucleotides, useful for
 PT diagnosing, treating and/or preventing cell proliferative, reproductive,
 PT gastrointestinal, neurological, urologic, and renal disorders.
 XX Claim 1; Page 166; 194pp; English.
 PS The invention relates to human organelle-associated proteins (ORGA) and
 CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
 CC for diagnosing, treating and/or preventing cell proliferative disorders
 CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
 CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
 CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
 CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
 CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
 CC disease, Pick's disease, Huntington's disease or dementia), urologic
 CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
 CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
 CC amyloidosis, renal failure, Addison's disease or hypertension). The
 CC invention is also useful in gene therapy. The present sequence is human
 CC ORGA protein
 XX Sequence 162 AA;
 SQ
 Query Match 54.0%; Score 134; DB 6; Length 162;
 Best Local Similarity 100.0%; Pred. No. 4.8e-127;
 Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNINA 60

CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
 CC amyloidosis, renal failure, Addison's disease or hypertension). The
 CC invention is also useful in gene therapy. The present sequence is human
 CC ORGA protein
 XX
 SQ Sequence 183 AA;
 Query Match 49.6%; Score 123; DB 6; Length 183;
 Best Local Similarity 100.0%; Pred. No. 7.4e-116;
 Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQTCPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVSVSGENTVMSCNISNA 60
 DB 1 MQTCPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVSVSGENTVMSCNISNA 60
 QY 61 FSHVNIKLRAHQESAIENEVAPGYFSDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAIENEVAPGYFSDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120
 QY 121 GHQ 123
 DB 121 GHQ 123
 RESULT 9
 AAO29895
 ID AAO29895 standard; protein; 101 AA.
 XX
 AC AAO29895;
 XX
 DT 03-SEP-2003 (first entry)
 XX
 DE Human organelle-associated protein (ORGA)-8.
 XX
 KW Human; organelle-associated protein; ORGA; cell proliferative disorder;
 KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
 KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
 KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
 KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
 KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
 KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
 KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
 KW infection; urethritis; prostatitis.
 XX
 OS Homo sapiens.
 XX
 PN WO2003044171-A2.
 XX
 PD 30-MAY-2003.
 XX
 PF 15-NOV-2002; 2002WO-US036807.
 XX
 PR 16-NOV-2001; 2001US-0332384P.
 PR 13-DEC-2001; 2001US-0341187P.
 PR 23-JAN-2002; 2002US-0351151P.
 PR 27-FEB-2002; 2002US-0360269P.
 PR 05-APR-2002; 2002US-0370637P.
 PR 14-JUN-2002; 2002US-0388946P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Rankumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
 PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
 PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
 PI Hafalia AJA, Bulloch S;
 XX
 DR WPI; 2003-457603/43.
 DR N-PSDB; AAL60548.
 XX
 PT New organelle-associated proteins and polynucleotides, useful for
 PT diagnosing, treating and/or preventing cell proliferative, reproductive,
 PT gastrointestinal, neurological, urologic, and renal disorders.
 XX

DB 1 MQTCPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVSVSGENTVMSCNISNA 60
 QY 61 FSHVNIKLRAHQESAIENEVAPGYFSDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAIENEVAPGYFSDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120
 QY 121 GHQNNRQVTLVS 134
 DB 121 GHQNNRQVTLVS 134
 RESULT 8
 AAO29897
 ID AAO29897 standard; protein; 183 AA.
 XX
 AC AAO29897;
 XX
 DT 03-SEP-2003 (first entry)
 XX
 DE Human organelle-associated protein (ORGA)-10.
 XX
 KW Human; organelle-associated protein; ORGA; cell proliferative disorder;
 KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
 KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
 KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
 KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
 KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
 KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
 KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
 KW infection; urethritis; prostatitis.
 XX
 OS Homo sapiens.
 XX
 PN WO2003044171-A2.
 XX
 PD 30-MAY-2003.
 XX
 PF 15-NOV-2002; 2002WO-US036807.
 XX
 PR 16-NOV-2001; 2001US-0332384P.
 PR 13-DEC-2001; 2001US-0341187P.
 PR 23-JAN-2002; 2002US-0351151P.
 PR 27-FEB-2002; 2002US-0360269P.
 PR 05-APR-2002; 2002US-0370637P.
 PR 14-JUN-2002; 2002US-0388946P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Rankumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
 PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
 PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
 PI Hafalia AJA, Bulloch S;
 XX
 DR WPI; 2003-457603/43.
 DR N-PSDB; AAL60550.
 XX
 PT New organelle-associated proteins and polynucleotides, useful for
 PT diagnosing, treating and/or preventing cell proliferative, reproductive,
 PT gastrointestinal, neurological, urologic, and renal disorders.
 XX
 PS Claim 1; Page 166-167; 194pp; English.
 XX
 CC The invention relates to human organelle-associated proteins (ORGA) and
 CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
 CC for diagnosing, treating and/or preventing cell proliferative disorders
 CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
 CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
 CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
 CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
 CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
 CC disease, Pick's disease, Huntington's disease or dementia), urologic
 CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper

PS Claim 1; Page 165-166; 194pp; English.

XX The invention relates to human organelle-associated proteins (ORGA) and
CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
CC for diagnosing, treating and/or preventing cell proliferative disorders
CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
CC disease, Pick's disease, Huntington's disease or dementia), urologic
CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
CC amyloidosis, renal failure, Addison's disease or hypertension). The
CC invention is also useful in gene therapy. The present sequence is human
XX ORGA protein

XX Sequence 101 AA;

Query Match 37.1%; Score 92; DB 6; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.1e-84;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQTCLAPFGHVSQALGTLFLFLASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
DB 1 MQTCLAPFGHVSQALGTLFLFLASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
QY 61 FSHVNIKLRAHQESAIFNEVAPGVFSRDGQ 92
DB 61 FSHVNIKLRAHQESAIFNEVAPGVFSRDGQ 92

RESULT 10
AAB34671
ID AAB34671 standard; protein; 55 AA.

XX AAB34671;

XX 26-JAN-2001 (first entry)

XX Gene 35 human secreted protein homologous amino acid sequence #155.

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW anti-rheumatic; antiproliferative; cytostatic; cardiac; vasotropic;
KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;
KW cerebrovascular disorder; angiogenesis; nervous system disorder;
KW ocular disorder; wound healing; skin aging; food additive; preservative.

XX Homo sapiens.

XX WO200056751-A1.

XX 28-SEP-2000.

XX 09-MAR-2000; 2000WO-US006013.

XX 19-MAR-1999; 93US-0125360P.

XX 11-JUN-1999; 93US-0138626P.

XX 03-DEC-1999; 93US-0168662P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI, 2000-579482/54.

XX Isolated nucleic acid molecule encoding a human secreted protein is used
XX in preventing, treating or ameliorating a medical condition.

XX Disclosure; Page 412; 419pp; English.

XX

CC The polynucleotide sequences given in AAC59738 to AAC59787 encode the
CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
CC AAB34686 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Example of activities include:
CC antiarthritic; immunosuppressive; anti-rheumatic; antiproliferative;
CC cytostatic; cardiac; vasotropic; cerebroprotective; neurotropic;
CC neuroprotective; antibacterial; virucide; fungicide; and
CC ophthalmological. The polynucleotides and proteins can be used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and
CC cancers of the breast or liver, cardiovascular disorders, cerebrovascular
CC disorders, angiogenesis, nervous system disorders, infections caused by
CC bacteria, viruses and fungi and ocular disorders. The proteins can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The proteins can also be used as a
CC food additive or preservative to increase or decrease storage
CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used
CC in the exemplification of the present invention

XX Sequence 55 AA;

Query Match 22.2%; Score 55; DB 3; Length 55;
Best Local Similarity 100.0%; Pred. No. 1.8e-47;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 LRAGAQQGLSRASAEELWTPDSEPTPRPLALVFKPSPLGALLSPQLPFPYAADP 248
DB 1 LRAGAQQGLSRASAEELWTPDSEPTPRPLALVFKPSPLGALLSPQLPFPYAADP 55

RESULT 11
AAO29898

ID AAO29898 standard; protein; 107 AA.

XX AAO29898;

XX 03-SEP-2003 (first entry)

XX Human organelle-associated protein (ORGA)-11.

XX Human; organelle-associated protein; ORGA; cell proliferative disorder;
KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
KW infection; urethritis; prostatitis.

XX Homo sapiens.

XX WO2003044171-A2.

XX 30-MAY-2003.

XX 15-NOV-2002; 2002WO-US036807.

XX 16-NOV-2001; 2001US-0332384P.

XX 13-DEC-2001; 2001US-0341187P.

XX 23-JAN-2002; 2002US-0351151P.

XX 27-FEB-2002; 2002US-0360269P.

XX 05-APR-2002; 2002US-0370637P.

XX 14-JUN-2002; 2002US-0388946P.

XX (INCY-) INCYTE GENOMICS INC.

XX Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
 PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
 PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
 PI Hafalia AJA, Bulloch S;
 XX WPI; 2003-457603/43.
 DR N-PSDB; AAL60551.
 XX New organelle-associated proteins and polynucleotides, useful for
 PT diagnosing, treating and/or preventing cell proliferative, reproductive,
 PT gastrointestinal, neurological, urologic, and renal disorders.
 XX Claim 1; Page 167; 194pp; English.
 XX The invention relates to human organelle-associated proteins (ORGA) and
 CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
 CC for diagnosing, treating and/or preventing cell proliferative disorders
 CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
 CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
 CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
 CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
 CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
 CC disease, Pick's disease, Huntington's disease or dementia), urologic
 CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
 CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
 CC amyloidosis, renal failure, Addison's disease or hypertension). The
 CC invention is also useful in gene therapy. The present sequence is human
 CC ORGA protein.
 XX Sequence 107 AA;
 SQ
 Query Match 22.2%; Score 55; DB 6; Length 107;
 Best Local Similarity 100.0%; Pred. No. 3.4e-47;
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQTCPLAFCHVSOALGTLFLAASLSAQNEGWDSPTCTGVVSVSGWNTWMC 55
 DB 1 MQTCPLAFCHVSOALGTLFLAASLSAQNEGWDSPTCTGVVSVSGWNTWMC 55
 RESULT 12
 AAB34672
 ID AAB34672 standard; protein; 55 AA.
 AC AAB34672;
 XX 26-JAN-2001 (first entry)
 XX Human secreted protein sequence encoded by gene 35 SEQ ID NO:156.
 XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
 KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
 KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
 KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
 KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;
 KW cerebrovascular disorder; angiogenesis; nervous system disorder;
 KW ocular disorder; wound healing; skin aging; food additive; preservative.
 XX Homo sapiens.
 OS WO200056751-A1.
 XX 28-SEP-2000.
 XX 09-MAR-2000; 2000WO-US006013.
 XX 19-MAR-1999; 99US-0125360P.
 PR 11-JUN-1999; 99US-0138626P.
 PR 03-DEC-1999; 99US-0168662P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 PI Rosen CA, Ruben SM, Komatsoulis G;
 XX WPI; 2000-579482/54.
 XX Isolated nucleic acid molecule encoding a human secreted protein is used
 PT in preventing, treating or ameliorating a medical condition.
 XX Disclosure; Page 412-413; 419pp; English.
 XX The polynucleotide sequences given in AAC59738 to AAC59787 encode the
 CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
 CC AAB34686 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Example of activities include:
 CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;
 CC cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic;
 CC neuroprotective; antibacterial; virucide; fungicide; and
 CC ophthalmological. The polynucleotides and proteins can be used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and
 CC cancers of the breast or liver, cardiovascular disorders, cerebrovascular
 CC disorders, angiogenesis, nervous system disorders, infections caused by
 CC bacteria, viruses and fungi and ocular disorders. The proteins can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The proteins can also be used as a
 CC food additive or preservative to increase or decrease storage
 CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used
 CC in the exemplification of the present invention
 XX Sequence 55 AA;
 SQ
 Query Match 19.4%; Score 48; DB 3; Length 55;
 Best Local Similarity 100.0%; Pred. No. 2.2e-40;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 194 LRAGAQGLSRASAEIWTPTDSEPTPRPLALVFKPSPLGALLESFQPL 241
 DB 1 LRAGAQGLSRASAEIWTPTDSEPTPRPLALVFKPSPLGALLESFQPL 48
 RESULT 13.
 ADB36337
 ID ADB36337 standard; protein; 43 AA.
 AC ADB36337;
 XX 04-DEC-2003 (first entry)
 XX Human immune response associated protein IRAP-8 SEQ ID NO:8.
 XX human; immune response associated protein; IRAP; anti-HIV; anti-allergic;
 KW antianemic; antisthmatic; antiarteriosclerotic; dermatological;
 KW antineoplastic; antidiabetic; nephroprotective; antithyroid; thyromimetic;
 KW immunosuppressive; antirheumatic; antiarthritic; osteopathic;
 KW antibacterial; virucide; antiparasitic; protozoacide; fungicide;
 KW cerebroprotective; neuroprotective; neurotropic; antiparkinsonian;
 KW antipsoriatic; cytostatic; cardiant; gene therapy;
 KW immune system disorder; neurological disorder; developmental disorder;
 KW muscle disorder; cell proliferative disorder; AIDS; allergy; anaemia;
 KW asthma; atherosclerosis; contact dermatitis; Crohn's disease; diabetes;
 KW glomerulonephritis; Grave's disease; Hashimoto's thyroiditis;
 KW multiple sclerosis; rheumatoid arthritis; osteoporosis;
 KW systemic lupus erythematosus; infection; stroke; Alzheimer's disease;
 KW Parkinson's disease; psoriasis; cancer; cardiomyopathy.
 XX Homo sapiens.
 OS

PN WO2003074726-A2.
 XX 12-SEP-2003.
 XX 28-FEB-2003; 2003WO-US006307.
 XX 01-MAR-2002; 2002US-0361088P.
 PR 27-MAR-2002; 2002US-0368494P.
 PR 10-MAY-2002; 2002US-0379876P.
 PR 28-JUN-2002; 2002US-0392641P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Chawla NK, Becha SD, Jin P, Tran UK, Kable AE, Lee S;
 PI Marquis JP, Emerling BM, Ison CH, Ramkumar J, Chang H;
 XX WPI; 2003-722079/68.
 DR N-PSDB; ADB36353.
 XX New human immune response associated proteins and polynucleotides, useful
 PT for diagnosing, preventing or treating diseases or conditions associated
 PT with aberrant enzyme expression, e.g. cancer, AIDS, atherosclerosis or
 PT stroke.
 XX Claim 1; Page 143; 158pp; English.
 XX ADB36346 to ADB36361 encode the human immune response associated proteins
 CC given in ADB36330 to ADB36345, designated IRAP-1 to IRAP-16. IRAP
 CC sequences have anti-HIV, anti-allergic, anti-inflammatory, antiasthmatic,
 CC antiarteriosclerotic, dermatological, antitumor, antidiabetic,
 CC nephrotropic, antithyroid, thyromimetic, immunosuppressive,
 CC antihemagic, antithyroid, osteoparatic, antibacterial, virucide,
 CC antiparasitic, protozoacide, fungicide, cerebroprotective,
 CC neuroprotective, neurotropic, antiparkinsonian, antipsoriatic, cytostatic
 CC and radiant activities, and can be used in gene therapy. The IRAP
 CC proteins and polynucleotides can be used in diagnosing, preventing or
 CC treating diseases or conditions associated with the decreased expression
 CC or overexpression of IRAP, such as immune system, neurological,
 CC developmental, muscle or cell proliferative disorders. The disorders may
 CC include AIDS, allergies, anaemia, asthma, atherosclerosis, contact
 CC dermatitis, Crohn's disease, diabetes, glomerulonephritis, Grave's
 CC disease, Hashimoto's thyroiditis, multiple sclerosis, rheumatoid
 CC arthritis, osteoporosis, systemic lupus erythematosus, infections (e.g.
 CC bacterial, viral, fungal, parasitic or protozoal), stroke, Alzheimer's
 CC disease, Parkinson's disease, psoriasis, cancer and cardiomyopathy. They
 CC are also useful in assessing the effects of exogenous compounds on the
 CC expression of nucleic acid and amino acid sequences of IRAP. The IRAP or
 CC its fragments are useful in screening compounds for effectiveness as
 CC agonist or antagonist of the polypeptides, or in altering the expression
 CC of the target polynucleotide and compounds that specifically bind to or
 CC modulate the activity of the polypeptide.
 XX Sequence 43 AA;
 XX Query Match 12.5%; Score 31; DB 7; Length 43;
 XX Best Local Similarity 100.0%; Pred. No. 2.9e-23;
 XX Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOTCLAPFGHVSQALGTLFLAASLSAQNE 31
 DB 1 MOTCLAPFGHVSQALGTLFLAASLSAQNE 31
 RESULT 14
 AAB34673
 ID AAB34673 standard; protein; 30 AA.
 XX AAB34673;
 XX AC
 XX 26-JAN-2001 (first entry)
 XX DT
 XX Gene 35 human secreted protein homologous amino acid sequence #157.
 XX DE
 XX AC

KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
 KW antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic;
 KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
 KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
 KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;
 KW cerebrovascular disorder; angiogenesis; nervous system disorder;
 KW ocular disorder; wound healing; skin aging; food additive; preservative.
 OS Homo sapiens.
 XX WO2000056751-A1.
 PN 28-SEP-2000.
 XX 09-MAR-2000; 2000WO-US006013.
 PD 19-MAR-1999; 99US-0125360P.
 PR 11-JUN-1999; 99US-0138626P.
 PR 03-DEC-1999; 99US-016862P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Ruben SM, Komatsoulis G;
 PI WPI; 2000-579482/54.
 DR Isolated nucleic acid molecule encoding a human secreted protein is used
 PT in preventing, treating or ameliorating a medical condition.
 PT Disclosure; Page 413; 419pp; English.
 XX The polynucleotide sequences given in AAC59738 to AAC59787 encode the
 CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
 CC AAB34686 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Example of activities include:
 CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;
 CC cytostatic; cardiac; vasotropic; cerebroprotective; neurotropic;
 CC neuroprotective; antibacterial; virucide; fungicide; and
 CC ophthalmological. The polynucleotides and proteins can be used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and
 CC cancers of the breast or liver, cardiovascular disorders, cerebrovascular
 CC disorders, angiogenesis, nervous system disorders, infections caused by
 CC bacteria, viruses and fungi and ocular disorders. The proteins can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The proteins can also be used as a
 CC food additive or preservative to increase or decrease storage
 CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used
 CC in the exemplification of the present invention
 XX Sequence 30 AA;
 XX Query Match 12.1%; Score 30; DB 3; Length 30;
 XX Best Local Similarity 100.0%; Pred. No. 2.1e-22;
 XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 62 SHVNIKLRAGQESAIIFNEVAPGYFSRDGW 91
 DB 1 SHVNIKLRAGQESAIIFNEVAPGYFSRDGW 30
 RESULT 15
 AAB34674
 ID AAB34674 standard; protein; 27 AA.
 XX AAB34674;
 XX AC

Search completed: March 20, 2005, 10:16:44
Job time : 91 secs

XX DT 26-JAN-2001 (first entry)
XX DE Human secreted protein sequence encoded by gene 35 SEQ ID NO:158.
XX DE Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW cerebroprotective; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;
KW cerebrovascular disorder; angiogenesis; nervous system disorder;
KW ocular disorder; wound healing; skin aging; food additive; preservative.
XX OS Homo sapiens.
XX PN WO200056751-A1.
XX PD 28-SEP-2000.
XX PF 09-MAR-2000; 2000WO-US006013.
XX PR 19-MAR-1999; 99US-0125360P.
XX PR 11-JUN-1999; 99US-0138626P.
XX PR 03-DEC-1999; 99US-0168662P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM, Komatsoulis G;
XX WPI; 2000-579482/54.
XX PT Isolated nucleic acid molecule encoding a human secreted protein is used
XX in preventing, treating or ameliorating a medical condition.
XX PS Disclosure; Page 413; 419pp; English.
XX CC The polynucleotide sequences given in AAC59738 to AAC59787 encode the
CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
CC AAB34686 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Example of activities include:
CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;
CC cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic;
CC neuroprotective; antibacterial; virucide; fungicide; and
CC ophthalmological. The polynucleotides and proteins can be used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and
CC cancers of the breast or liver, cardiovascular disorders, cerebrovascular
CC disorders, angiogenesis, nervous system disorders, infections caused by
CC bacteria, viruses and fungi and ocular disorders. The proteins can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The proteins can also be used as a
CC food additive or preservative to increase or decrease storage
CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used
CC in the exemplification of the present invention
XX SQ Sequence 27 AA;

Query Match 4.8%; Score 12; DB 3; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 62 SHVNIKLRHQGQ 73
Db 1 SHVNIKLRHQGQ 12

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 10:13:54 ; Search time 26 Seconds
(without alignments)
712.037 Million cell updates/sec

Title: US-10-080-522-1
Perfect score: 248
Sequence: 1 MOTCPAPFGHVSQLGTL.....FLGALELLSQPLFPYADP 248

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	248	100.0	248	US-08-755-559-1	Sequence 1, Appli
2	248	100.0	248	US-09-210-474-1	Sequence 1, Appli
3	248	100.0	248	US-09-539-774-1	Sequence 1, Appli
4	248	100.0	248	US-09-997-165-4	Sequence 4, Appli
5	248	100.0	248	US-09-949-016-6297	Sequence 6297, Ap
6	248	100.0	258	US-09-949-016-8807	Sequence 8807, Ap
7	8	3.2	453	US-09-252-991A-31634	Sequence 31634, A
8	7	2.8	67	US-09-248-796A-23640	Sequence 23640, A
9	7	2.8	81	US-09-621-976-5355	Sequence 5255, Ap
10	7	2.8	83	US-09-270-767-60208	Sequence 60208, A
11	7	2.8	93	US-09-902-540-13818	Sequence 13818, A
12	7	2.8	138	US-09-148-545-222	Sequence 222, App
13	7	2.8	139	US-09-148-545-160	Sequence 160, App
14	7	2.8	197	US-09-252-991A-32189	Sequence 32189, A
15	7	2.8	202	US-09-902-540-16034	Sequence 16034, A
16	7	2.8	210	US-10-006-011A-10	Sequence 10, Appl
17	7	2.8	231	US-09-583-110-3995	Sequence 3995, Ap
18	7	2.8	233	US-09-252-991A-21274	Sequence 21274, A
19	7	2.8	235	US-09-390-721-4	Sequence 4, Appli
20	7	2.8	235	US-09-489-039A-12572	Sequence 12572, A
21	7	2.8	235	US-09-713-893-4	Sequence 4, Appli
22	7	2.8	242	US-08-663-310-9	Sequence 9, Appli
23	7	2.8	242	US-09-006-491-9	Sequence 9, Appli
24	7	2.8	242	US-09-135-919-9	Sequence 9, Appli
25	7	2.8	242	US-08-980-832-29	Sequence 29, Appl
26	7	2.8	242	US-09-920-923B-29	Sequence 29, Appl
27	7	2.8	248	US-09-050-739-60	Sequence 60, Appl

ALIGNMENTS

RESULT 1

US-08-755-559-1
; Sequence 1, Application US/08755559
; Patent No. 5912142
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,559
; FILING DATE: 22-NOV-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-755-559-1

Query Match 100.0%; Score 248; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 9.8e-238;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCPAPFGHVSQLGTLFLAASLSAQNAGWDSPTCTEGVSVSWGENTVMSCNISNA 60
DB 1 MOTCPAPFGHVSQLGTLFLAASLSAQNAGWDSPTCTEGVSVSWGENTVMSCNISNA 60

us-10-080-522-1.oli.ra1

Sun Mar 20 14:17:41 2005

QY 61 FSHVNIKLAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
 QY 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180
 DB 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180
 QY 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWPDPSEPTPRPLALVFKPSPLGALLELLSPQ 240
 DB 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWPDPSEPTPRPLALVFKPSPLGALLELLSPQ 240
 QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 2
 US-09-210-474-1
 ; Sequence 1, Application US/09210474
 ; Patent No. 6072034
 ; GENERAL INFORMATION:
 ; APPLICANT: KAUFMAN, RUSSEL E.
 ; APPLICANT: SLENTZ-KESLER, KIMBERLY
 ; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
 ; TITLE OF INVENTION: CELLS
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHUYE P.C.
 ; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: U.S.A.
 ; ZIP: 22201-4714
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/210,474
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/755,559
 ; FILING DATE: 22-NOV-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WILSON, MARY J.
 ; REGISTRATION NUMBER: 32,955
 ; REFERENCE/DOCKET NUMBER: 1579-116
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 816-4000
 ; TELEFAX: (703) 816-4100
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 248 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-210-474-1

Query Match 100.0%; Score 248; DB 3; Length 248;
 Best Local Similarity 100.0%; Pred. No. 9.8e-238;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 DB 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 QY 61 FSHVNIKLAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120

QY 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180
 DB 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180
 QY 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWPDPSEPTPRPLALVFKPSPLGALLELLSPQ 240
 DB 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWPDPSEPTPRPLALVFKPSPLGALLELLSPQ 240
 QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 3
 US-09-539-774-1
 ; Sequence 1, Application US/09539774
 ; Patent No. 6350615
 ; GENERAL INFORMATION:
 ; APPLICANT: KAUFMAN, RUSSEL E.
 ; APPLICANT: SLENTZ-KESLER, KIMBERLY
 ; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
 ; TITLE OF INVENTION: CELLS
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHUYE P.C.
 ; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: U.S.A.
 ; ZIP: 22201-4714
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/539,774
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/210,474
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WILSON, MARY J.
 ; REGISTRATION NUMBER: 32,955
 ; REFERENCE/DOCKET NUMBER: 1579-116
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 816-4000
 ; TELEFAX: (703) 816-4100
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 248 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-539-774-1

Query Match 100.0%; Score 248; DB 3; Length 248;
 Best Local Similarity 100.0%; Pred. No. 9.8e-238;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 DB 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
 QY 61 FSHVNIKLAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLAHQESAI FNEVAPGYFSRDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
 QY 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180
 DB 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180

Db 121 GHQNNRQVTLVSGAEFQSPADTGFVPVAVTAVFILLVALVWFAYRCRCQQORREK 180
Qy 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALLELLSQP 240
Db 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALLELLSQP 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 4

US-09-997-165-4
; Sequence 4, Application US/09997165
; Patent No. 6762030
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fanslow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-997-165-4

Query Match 100.0%; Score 248; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 9.8e-238;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MOTCLAPFGHVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 60
Db 1 MOTCLAPFGHVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 60
Qy 61 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMMHLV 120
Db 61 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMMHLV 120
Qy 121 GHQNNRQVTLVSGAEFQSPADTGFVPVAVTAVFILLVALVWFAYRCRCQQORREK 180
Db 121 GHQNNRQVTLVSGAEFQSPADTGFVPVAVTAVFILLVALVWFAYRCRCQQORREK 180
Qy 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALLELLSQP 240
Db 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALLELLSQP 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 5

US-09-949-016-6297
; Sequence 6297, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6297
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6297

Query Match 100.0%; Score 248; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 9.8e-238;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MOTCLAPFGHVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 60
Db 1 MOTCLAPFGHVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 60
Qy 61 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMMHLV 120
Db 61 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMMHLV 120
Qy 121 GHQNNRQVTLVSGAEFQSPADTGFVPVAVTAVFILLVALVWFAYRCRCQQORREK 180
Db 121 GHQNNRQVTLVSGAEFQSPADTGFVPVAVTAVFILLVALVWFAYRCRCQQORREK 180
Qy 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALLELLSQP 240
Db 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPLGALLELLSQP 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 6

US-09-949-016-8807
; Sequence 8807, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8807
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8807

Query Match 100.0%; Score 248; DB 4; Length 258;
Best Local Similarity 100.0%; Pred. No. 1e-237;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MOTCLAPFGHVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 60
Db 11 MOTCLAPFGHVHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCNISNA 70
Qy 61 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMMHLV 120
Db 71 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMMHLV 130

QY 121 GHORNRQVTLVSGAEPSAPDTGFWPVAVTAVFILLVALVMFAWYRCRCSCQRRREK 180
 DB 131 GHORNRQVTLVSGAEPSAPDTGFWPVAVTAVFILLVALVMFAWYRCRCSCQRRREK 190
 QY 181 KFFLEPQKVAALRAGAQOGLSRASABELWTPDSETPRPLALVFKPSPLGALELLSQP 240
 DB 191 KFFLEPQKVAALRAGAQOGLSRASABELWTPDSETPRPLALVFKPSPLGALELLSQP 250
 QY 241 LPFYAADP 248
 DB 251 LPFYAADP 258

RESULT 7
 US-09-252-991A-31634
 ; Sequence 31634, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,789
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 31634
 ; LENGTH: 453
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-31634

Query Match 3.2%; Score 8; DB 4; Length 453;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 QGLSRASA 207
 DB 419 QGLSRASA 426

RESULT 8
 US-09-248-796A-23640
 ; Sequence 23640, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; PRIOR FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 23640
 ; LENGTH: 67
 ; TYPE: PRT
 ; ORGANISM: Candida albicans
 US-09-248-796A-23640

Query Match 2.8%; Score 7; DB 4; Length 67;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 QORREKK 181
 DB 37 QORREKK 43

RESULT 9
 US-09-621-976-5255
 ; Sequence 5255, Application US/09621976
 ; Patent No. 6639063
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Jobert, S.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
 ; FILE REFERENCE: GENSET.054PR2
 ; CURRENT APPLICATION NUMBER: US/09/621,976
 ; CURRENT FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 19335
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 5255
 ; LENGTH: 81
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SIGNAL
 ; LOCATION: -15...-1
 ; NAME/KEY: UNSURE
 ; LOCATION: 62
 ; OTHER INFORMATION: Xaa = *, Gln
 ; NAME/KEY: UNSURE
 ; LOCATION: 54
 ; OTHER INFORMATION: Xaa = Asp, Phe, Ile, Asn, Val, Tyr
 ; NAME/KEY: UNSURE
 ; LOCATION: 10
 ; OTHER INFORMATION: Xaa = Asp, Tyr
 ; NAME/KEY: UNSURE
 ; LOCATION: 14
 ; OTHER INFORMATION: Xaa = Cys, Arg
 ; NAME/KEY: UNSURE
 ; LOCATION: 64
 ; OTHER INFORMATION: Xaa = Cys, Gly
 ; NAME/KEY: UNSURE
 ; LOCATION: 38
 ; OTHER INFORMATION: Xaa = His, Tyr
 ; NAME/KEY: UNSURE
 ; LOCATION: 52
 ; OTHER INFORMATION: Xaa = Lys, Gln
 ; NAME/KEY: UNSURE
 ; LOCATION: 11
 ; OTHER INFORMATION: Xaa = Phe, Ile
 US-09-621-976-5255

Query Match 2.8%; Score 7; DB 4; Length 81;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 TLLFLAA 24
 DB 5 TLLFLAA 11

RESULT 10
 US-09-270-767-60208
 ; Sequence 60208, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 60208
 ; LENGTH: 83
 ; TYPE: PRT

```
; ORGANISM: Drosophila melanogaster
US-09-270-767-60208

Query Match          2.8%; Score 7; DB 4; Length 83;
Best Local Similarity 100.08; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 VPAVVTA 155
   |||||
Db 64 VPAVVTA 70

RESULT 11
US-09-902-540-13818
; Sequence 13818, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13818
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13818

Query Match          2.8%; Score 7; DB 4; Length 93;
Best Local Similarity 100.08; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 SLSAQNE 31
   |||||
Db 48 SLSAQNE 54

RESULT 12
US-09-148-545-222
; Sequence 222, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
```

us-10-080-522-1.oli.ra1

Sun Mar 20 14:17:41 2005

```
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 05-Sep-1997
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 222
; LENGTH: 138

Query Match      2.8%  Score 7;  DB 4;  Length 138;
Best Local Similarity 100.0%;  Pred. No. 52;
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

Oy      18  TLLFLAA 24
Db       5  TLLFLAA 11
|||||
|||||

RESULT 13
US-09-148-545-160
; Sequence 160, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PZ001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
```


[illegible]

Sun Mar 20 14:17:41 2005

```
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 160
; LENGTH: 139

Query Match      2.8%; Score 7; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 TLLFLAA 24
Db 5 TLLFLAA 11

RESULT 14
US-09-252-991A-32189
; Sequence 32189, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196, 116
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32189
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32189

Query Match      2.8%; Score 7; DB 4; Length 197;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 GLSRASA 207
Db 133 GLSRASA 139

RESULT 15
US-09-902-540-16034
; Sequence 16034, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
```

```
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16034
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-16034
```

```
Query Match      2.8%; Score 7; DB 4; Length 202;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 22 LAASLSA 28
Db 12 LAASLSA 18
```

```
Search completed: March 20, 2005, 10:19:23
Job time : 28 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 10:18:25 ; Search time 73 Seconds
(without alignments)
1122.934 Million cell updates/sec

Title: US-10-080-522-1

Perfect score: 248

Sequence: 1 MQTCPLAPGHVSQALGTL.....PLGALLELLSPQLPPYAADP 248

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1401741 seqs, 330541175 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	248	100.0	248	9	US-09-799-777-32
2	248	100.0	248	9	US-09-997-165-4
3	248	100.0	248	14	US-10-080-522-1
4	9	3.6	1709	14	US-10-156-761-11180
5	8	3.2	75	16	US-10-437-963-110606
6	8	3.2	123	16	US-10-437-963-116750
7	8	3.2	125	15	US-10-424-599-182088
8	8	3.2	294	15	US-10-282-122A-48538
9	8	3.2	384	15	US-10-282-122A-76494
10	8	3.2	545	14	US-10-156-761-11660
11	8	3.2	691	15	US-10-369-493-10173
12	7	2.8	19	15	US-10-468-372-9
13	7	2.8	36	11	US-09-833-245-858

Sequence 29093, A
Sequence 50684, A
Sequence 2378, Ap
Sequence 195864,
Sequence 145761,
Sequence 134138,
Sequence 342, App
Sequence 170987,
Sequence 9429, Ap
Sequence 3112, Ap
Sequence 2594, Ap
Sequence 176, App
Sequence 252, App
Sequence 222, App
Sequence 222, App
Sequence 4, Appli
Sequence 160, App
Sequence 39979, A
Sequence 165259,
Sequence 2088, Ap
Sequence 2, Appli
Sequence 1399, Ap
Sequence 2323, Ap
Sequence 184172,
Sequence 155, App
Sequence 155, App
Sequence 175211,
Sequence 175849,
Sequence 132222,

14 56 14 US-10-029-386-29093
15 58 16 US-10-767-701-50684
16 57 11 US-09-864-408A-2978
17 72 16 US-10-437-963-195864
18 78 15 US-10-424-599-145761
19 85 16 US-10-437-963-124138
20 88 14 US-10-195-730-342
21 88 16 US-10-799-747-342
22 91 16 US-10-437-963-170987
23 7 2.8 115 14 US-10-156-761-9429
24 7 2.8 128 15 US-10-104-047-3112
25 7 2.8 128 15 US-10-108-260A-2594
26 7 2.8 134 9 US-09-881-752A-176
27 7 2.8 138 9 US-09-731-872-252
28 7 2.8 138 9 US-09-981-876-222
29 7 2.8 138 10 US-09-148-545-222
30 7 2.8 138 10 US-09-876-997-252
31 7 2.8 138 15 US-10-468-372-4
32 7 2.8 139 9 US-09-981-876-160
33 7 2.8 139 10 US-09-148-545-160
34 7 2.8 142 16 US-10-767-701-39979
35 7 2.8 144 16 US-10-437-963-165259
36 7 2.8 147 15 US-10-264-237-2088
37 7 2.8 154 10 US-09-732-436-2
38 7 2.8 155 15 US-10-276-774-1399
39 7 2.8 160 15 US-10-094-749-2923
40 7 2.8 171 16 US-10-437-963-184172
41 7 2.8 174 14 US-10-097-065-155
42 7 2.8 174 15 US-10-372-876-155
43 7 2.8 177 16 US-10-437-963-175211
44 7 2.8 178 15 US-10-424-599-175849
45 7 2.8 187 16 US-10-437-963-132222

ALIGNMENTS

RESULT 1

US-09-799-777-32

; Sequence 32, Application US/09799777

; Patent No. US20020091244A1

; GENERAL INFORMATION:

; APPLICANT: Lal, Preeti

; Hillman, Jennifer L.

; Corley, Neil C.

; Guegler, Karl J.

; Baugh, Mariah

; Sather, Susan

; Shah, Purvi

; TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS

; NUMBER OF SEQUENCES: 154

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/799,777

; FILING DATE: 06-Mar-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/002,485

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: BILLINGS, LUCY J.

; REGISTRATION NUMBER: 36,749

us-10-080-522-1-oli.rapb

Sun Mar 20 14:17:41 2005

```

; REFERENCE/DOCKET NUMBER: PF-0459 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 32 :
US-09-799-777-32

Query Match      100.0%; Score 248; DB 9; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.8e-230; Indels 0; Gaps 0;
Matches 248; Conservative 0; Mismatches 0;

QY 1 MOTCPFLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
DB 1 MOTCPFLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120

QY 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAYRCRCQORREK 180
DB 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVMFAYRCRCQORREK 180

QY 181 KFFLLPQMKVAALRAGAQQGLSRASAEIWTDPSETPRPLALVFKPSPGLGALELLSQP 240
DB 181 KFFLLPQMKVAALRAGAQQGLSRASAEIWTDPSETPRPLALVFKPSPGLGALELLSQP 240

QY 241 LFPYAADP 248
DB 241 LFPYAADP 248

RESULT 2
US-09-997-165-4
; Sequence 4, Application US/09997165
; Patent No. US20020141999A1
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fanslow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-997-165-4

Query Match      100.0%; Score 248; DB 9; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.8e-230; Indels 0; Gaps 0;
Matches 248; Conservative 0; Mismatches 0;

QY 1 MOTCPFLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
DB 1 MOTCPFLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120

; REFERENCE/DOCKET NUMBER: PF-0459 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 32 :
US-10-080-522-1

Query Match      100.0%; Score 248; DB 14; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.8e-230; Indels 0; Gaps 0;
Matches 248; Conservative 0; Mismatches 0;

QY 1 MOTCPFLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
DB 1 MOTCPFLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120

; REFERENCE/DOCKET NUMBER: PF-0459 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 32 :
US-10-080-522-1

Query Match      100.0%; Score 248; DB 14; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.8e-230; Indels 0; Gaps 0;
Matches 248; Conservative 0; Mismatches 0;

QY 1 MOTCPFLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
DB 1 MOTCPFLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLYMHLV 120

```

Db 61 FSHVNIKURAHQGEISAFNEVAPGYFSRDRGWLQVGGVQAQLVIRKARDSHAGLYMWHLV 120
QY 121 GQRNNRQVTLVSGAEPOSADPTGFWPVPVAVTVAVFILLVALVMFAWYRCRCQORREK 180
Db 121 GQRNNRQVTLVSGAEPOSADPTGFWPVPVAVTVAVFILLVALVMFAWYRCRCQORREK 180
QY 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVPKPSPLGALELLSQP 240
Db 181 KFFLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVPKPSPLGALELLSQP 240
QY 241 LPFYAADP 248
Db 241 LPFYAADP 248

RESULT 4
US-10-156-761-11180
; Sequence 11180, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11180
; LENGTH: 1709
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11180

Query Match 3.6%; Score 9; DB 14; Length 1709;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 229 PLGALELLS 237
Db 1063 PLGALELLS 1071

RESULT 5
US-10-437-963-110606
; Sequence 110606, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 110606
; LENGTH: 75
; TYPE: PRT

; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_14657C.1.pap
US-10-437-963-110606

Query Match 3.2%; Score 8; DB 16; Length 75;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 VTAVFILL 160
Db 58 VTAVFILL 65

RESULT 6
US-10-437-963-116750
; Sequence 116750, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 116750
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(123)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_20220C.1.pap
US-10-437-963-116750

Query Match 3.2%; Score 8; DB 16; Length 123;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 LGALELLS 237
Db 21 LGALELLS 28

RESULT 7
US-10-424-599-182088
; Sequence 182088, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 182088
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Glycine max

APPLICANT: HATTORI, MASAIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11660
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11660

Query Match 3.2%; Score 8; DB 14; Length 545;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SPLGALEL 235
Db 120 SPLGALEL 127

RESULT 11
US-10-369-493-10173
; Sequence 10173, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Cher, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10173
; LENGTH: 691
; TYPE: PRT
; ORGANISM: magnetite-containing magnetic coccus
US-10-369-493-10173

Query Match 3.2%; Score 8; DB 15; Length 691;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LGTLLFLA 23
Db 7 LGTLLFLA 14

RESULT 12
US-10-468-372-9
; Sequence 9, Application US/10468372
; Publication No. US20040081984A1
; GENERAL INFORMATION:
; APPLICANT: The University of Virginia Patent Foundation
; APPLICANT: Laurie, Gordon W
; APPLICANT: Kumar, Rajesh W
; APPLICANT: Sanghi, Sandhya
; APPLICANT: Lumsden, Angela
; TITLE OF INVENTION: Ocular Tear Growth Factor-Like Protein
; FILE REFERENCE: 00662-02
; CURRENT APPLICATION NUMBER: US/10/468,372
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US 60/269,900

; PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-468-372-9

Query Match 2.8%; Score 7; DB 15; Length 19;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 TLLFLAA 24
Db 5 TLLFLAA 11

RESULT 13
US-09-833-245-858
; Sequence 858, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 858
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (17)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-833-245-858

Query Match 2.8%; Score 7; DB 11; Length 36;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LLFLAAS 25
Db 18 LLFLAAS 24

RESULT 14
US-10-029-386-29093
; Sequence 29093, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29093
; LENGTH: 56
; TYPE: PRT

Sun Mar 20 14:17:41 2005

```

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
US-10-029-386-29093

```

```

Query Match      2.8%  Score 7;  DB 14;  Length 56;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 133 VSGAEPQ 139
    |||||
DB 35 VSGAEPQ 41

```

```

RESULT 15
US-10-767-701-50684
; Sequence 50684, Application US/10767701
; Publication No. US20040172884A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 50684
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3480-024-PI-K1-H1.pep
US-10-767-701-50684

```

```

Query Match      2.8%  Score 7;  DB 16;  Length 58;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 181 KFFLLEP 187
    |||||
DB 15 KFFLLEP 21

```

Search completed: March 20, 2005, 10:25:26
Job time : 76 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 10:13:31 ; Search time 25 Seconds
(without alignments)
954.470 Million cell updates/sec

Title: US-10-080-522-1
Perfect score: 248
Sequence: 1 MQTCPLAPGHVSQALGTL.....PLGALELLSQPLFPYAADP 248

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB IP	Description
1	8	3.2	108	2 F72549	hypothetical prote
2	8	3.2	384	2 E71331	probable cell divi
3	8	3.2	10797	2 T30192	probable peptide s
4	7	2.8	51	2 C82707	hypothetical prote
5	7	2.8	74	2 T16336	hypothetical prote
6	7	2.8	96	2 T22979	hypothetical prote
7	7	2.8	131	2 AE0126	conserved hypothet
8	7	2.8	131	4 S63446	hypothetical prote
9	7	2.8	134	2 D64558	hypothetical prote
10	7	2.8	136	2 S37924	hypothetical prote
11	7	2.8	143	2 A97133	uncharacterized ph
12	7	2.8	156	2 F82575	DNA methylase Xf22
13	7	2.8	162	2 S70593	NADH2 dehydrogenas
14	7	2.8	180	1 JQ0986	lipid body-associat
15	7	2.8	181	2 AE0435	conserved hypothet
16	7	2.8	181	2 B82134	transcription regu
17	7	2.8	195	2 A99346	hypothetical prote
18	7	2.8	207	2 B44994	eggshell protein 2
19	7	2.8	212	2 A44994	eggshell protein 1
20	7	2.8	225	2 T25075	hypothetical prote
21	7	2.8	229	2 F70568	hypothetical prote
22	7	2.8	234	2 C81149	hypothetical prote
23	7	2.8	237	2 G46678	probable RING zinc
24	7	2.8	243	2 B82574	site-specific DNA-
25	7	2.8	246	2 G83959	tRNA methyltransfe
26	7	2.8	248	2 H70511	probable prCA prot
27	7	2.8	259	2 F82087	deoxyribose-phosph
28	7	2.8	265	2 E87074	proteasome [alpha]
29	7	2.8	265	2 S72865	hypothetical prote

30 7 2.8 272 2 A69961 geranyltranstranf
31 7 2.8 276 2 E96951 conserved membrane
32 7 2.8 284 1 E64938 probable aldehyde
33 7 2.8 284 2 F85788 probable an aldehy
34 7 2.8 284 2 F80940 probable an aldehy
35 7 2.8 288 1 S70955 ctnr protein - Vib
36 7 2.8 288 2 T44314 hypothetrical prote
37 7 2.8 288 2 AE2947 ABC transporter, m
38 7 2.8 290 2 G75470 conserved hypothet
39 7 2.8 290 2 AE1126 hypothetrical prote
40 7 2.8 299 2 B70737 hypothetrical prote
41 7 2.8 304 2 B92244 probable permease
42 7 2.8 304 2 AE3062 hypothetrical prote
43 7 2.8 307 2 G82385 hypothetrical prote
44 7 2.8 311 2 E98335 membrane protein n
45 7 2.8 312 2 AD2797 malonate transport

ALIGNMENTS

RESULT 1

F72549
hypothetical protein APE1682 - Aeropyrum pernix (strain Kl)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C:Accession: F72549

R:Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: F72549

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-108 <KAW>

A:Cross-references: UNIPROT:Q9YBB5; DBJ:AP000062; NID:G5105244; PIDN:BAA80683.1; PID:G5

A:Experimental source: strain Kl

C:Genetics:

A:Gene: APE1682

C:Superfamily: Aeropyrum pernix hypothetical protein APE1682

Query Match 3.2%; Score 8; DB 2; Length 108;

Best Local Similarity 100.0%; Pred. No. 2.6;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 217 TPRPLALV 224

Db 19 TPRPLALV 26

RESULT 2

E71331

probable cell division protein (ftsW) - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C:Accession: E71331

R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin

son, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McD

they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998

A>Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71250; MUID:98332770; PMID:9665876

A:Accession: E71331

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-384 <COL>

A:Cross-references: UNIPROT:O83402; GB:AE001217; GB:AE000520; NID:G3322656; PIDN:AAC6537

A:Experimental source: strain Nichols

C:Genetics:

A:Gene: TP0387

C:Superfamily: rod shape-determining protein

Query Match 3.2%; Score 8; DB 2; Length 384;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 VFILVAL 163
19 VFILVAL 26
|||||

Db

RESULT 3
T30192
probable peptide synthetase - Aureobasidium pullulans
C:Species: Aureobasidium pullulans
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 09-Jul-2004
C:Accession: T30192
R:Peery, R.B.; Thornwell, S.J.; Tobin, M.B.; Skatrud, P.L.
submitted to the EMBL Data Library, January 1997
A:Description: Discovery of an MDR-like gene adjacent to a peptide synthetase in Aureoba
A:Reference number: Z20767
A:Accession: T30192
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-10797 <PE>
A:Cross-references: UNIPROT:Q94116; EMBL:U85909; NID:G4099310; PID:G4099313; PIDN:AA0005
C:Genetics:
A:Introns: 2078/3; 4142/3; 6075/3; 7963/3; 7985/2
C:Keywords: carrier protein
F:1618-1688/Domain: acyl carrier protein homology <ACP1>
F:1618-1688/Domain: acyl carrier protein homology <ACP2>
F:3682-3752/Domain: acyl carrier protein homology <ACP3>
F:5615-5685/Domain: acyl carrier protein homology <ACP4>
F:7503-7573/Domain: acyl carrier protein homology <ACP4>
F:9683-9752/Domain: acyl carrier protein homology <ACP5>

Query Match 3.2%; Score 8; DB 2; Length 10797;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 LGALELLS 237
598 LGALELLS 605
|||||

Db

RESULT 4
C82707
hypothetical protein XF1235 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: C82707
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: C82707
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-51 <SIM>
A:Cross-references: UNIPROT:Q9PDZ3; GB:AE003957; GB:AE003849; NID:G9106207; PIDN:AAF8404
R:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Prohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation

C:Genetics:
A:Gene: XF1235

Query Match 2.8%; Score 7; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 VAALRAG 197
12 VAALRAG 18
|||||

Db

RESULT 5
T16336
hypothetical protein F41G3.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16336
R:Manezes, S.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F41G3.
A:Reference number: Z18496
A:Accession: T16336
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-74 <MEN>
A:Cross-references: UNIPROT:Q20295; EMBL:U39999; NID:G1055133; PID:G1055142; PIDN:AAA811
C:Genetics:
A:Gene: CESP:F41G3.1
A:Introns: 22/1

Query Match 2.8%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 GLSRASA 207
61 GLSRASA 67
|||||

Db

RESULT 6
T22979
hypothetical protein F59B10.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22979
R:Lloyd, C.
submitted to the EMBL Data Library, March 1995
A:Reference number: Z19646
A:Accession: T22979
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-96 <WIL>
A:Cross-references: UNIPROT:Q09953; EMBL:Z48716; PIDN:CAA88599.1; GSPDB:GN000020; CESP:F
A:Experimental source: clone F59B10
C:Genetics:
A:Gene: CESP:F59B10.5
A:Map position: 2
A:Introns: 33/3; 73/2

Query Match 2.8%; Score 7; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 FILLVAL 163
4 FILLVAL 10
|||||

Db

RESULT 7
AE0126
conserved hypothetical protein YPO1030 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
 C/Accession: AE0126
 R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell, Nature 413, 523-527, 2001
 A>Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A/Reference number: AB0001; MUID:21470413; PMID:11586360
 A/Accession: AE0126
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-131 <CUR>
 A/Cross-references: UNIPROT:Q8ZH79; GB:AL590842; PIDN:CAC99872.1; PID:gl5979097; GSPDB:C
 C/Genetics:
 A:Gene: YPO1030

Query Match 2.8%; Score 7; DB 2; Length 131;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 ALGTLFL 21
 |||||
 Db 77 ALGTLFL 83

RESULT 8
 S63446
 hypothetical protein cosmid PEOA461 - yeast (*Saccharomyces cerevisiae*)
 C/Species: *Saccharomyces cerevisiae*
 C/Date: 16-May-1996 #sequence_revision 05-Sep-1996 #text_change 09-Jul-2004
 C/Accession: S71984; S63446
 R/Casamayor, A.; Khalid, H.; Baicells, L.; Aldea, M.; Casas, C.; Herrero, E.; Arino, J. Yeast 12, 1013-1020, 1996
 A>Title: Sequence analysis of a 13.4 kbp fragment from the left arm of chromosome XV revealing frames.
 A/Reference number: S71978; MUID:97051588; PMID:8896265
 A/Accession: S71984
 A/Status: nucleic acid sequence not shown; translation not shown; conceptual translation
 A/Molecule type: DNA
 A/Residues: 1-131 <CAS>
 A/Cross-references: UNIPROT:Q02812; EMBL:U01293; NID:gl209710; PIDN:AAC49468.1; PID:gl209710
 A/Experimental source: strain F11679
 A/Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1995
 C/Comment: There is no evidence that this sequence is expressed.
 C/Genetics:
 A:Map position: 15L
 C/Keywords: pseudogene

Query Match 2.8%; Score 7; DB 4; Length 131;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 157 FILLVAL 163
 |||||
 Db 97 FILLVAL 103

RESULT 9
 D64558
 hypothetical protein HP0308 - *Helicobacter pylori* (strain 26695)
 C/Species: *Helicobacter pylori*
 C/Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
 C/Accession: D64558
 R/Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997
 A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
 A>Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
 A/Reference number: A64520; MUID:97394467; PMID:9252185
 A/Accession: D64558
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA

A/Residues: 1-134 <TOM>
 A/Cross-references: UNIPROT:O25078; GB:AE000549; PIDN:AA81936.1; PID:g4861

Query Match 2.8%; Score 7; DB 2; Length 134;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LGTLLFL 22
 |||||
 Db 97 LGTLLFL 103

RESULT 10
 S37924
 hypothetical protein YKL097c - yeast (*Saccharomyces cerevisiae*)
 A/Alternate names: hypothetical protein YKL445
 C/Species: *Saccharomyces cerevisiae*
 C/Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
 C/Accession: S37924; S39088
 R/Cheret, G.; Fukuhara, H.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Pallier, C.; Puzo submitted to the Protein Sequence Database, March 1994
 A/Reference number: S37920
 A/Accession: S37924
 A/Molecule type: DNA
 A/Residues: 1-136 <CHR>
 A/Cross-references: UNIPROT:P34245; EMBL:Z28097; NID:g486160; PIDN:CAA81936.1; PID:g4861
 A/Experimental source: strain S288C
 R/Pallier, C.; Valens, M.; Puzos, V.; Fukuhara, H.; Cheret, G.; Sor, F.; Bolotin-Fukuhara Yeast 9, 1149-1155, 1993
 A>Title: DNA sequence analysis of a 17 kb fragment of yeast chromosome XI physically located in protein kinases.
 A/Reference number: S39084; MUID:94078677; PMID:8256524
 A/Accession: S39088
 A/Status: translation not shown
 A/Molecule type: DNA
 A/Residues: 1-136 <PAL>
 A/Cross-references: EMBL:X71133; NID:g431205; PIDN:CAA50460.1; PID:g431219
 A/Experimental source: strain S288C
 C/Genetics:
 A/Cross-references: SGD:S0001580
 A/Map position: 11L
 C/Superfamily: *Saccharomyces* hypothetical protein YKL097c

Query Match 2.8%; Score 7; DB 2; Length 136;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 147 WPVPAVV 153
 |||||
 Db 17 WPVPAVV 23

RESULT 11
 A97133
 uncharacterized phage related protein [imported] - *Clostridium acetobutylicum*
 C/Species: *Clostridium acetobutylicum*
 C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
 C/Accession: A97133
 R/Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R. J. Bacteriol. 183, 4823-4838, 2001
 A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium acetobutylicum* A96900; MUID:21359325; PMID:21359325
 A/Reference number: A96900; MUID:21359325; PMID:21359325
 A/Accession: A97133
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-143 <CUR>
 A/Cross-references: UNIPROT:Q97HW8; GB:AB001437; PIDN:AAK79852.1; PID:gl5024868; GSPDB:G
 A/Experimental source: *Clostridium acetobutylicum* ATCC824
 C/Genetics:
 A:Gene: CAC1888

Query Match 2.8%; Score 7; DB 2; Length 143;

Query Match 2.8%; Score 7; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 ALRGAQ 199
| | | | |
Db 31 ALRGAQ 37

RESULT 12
DNA methylase XP2297 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: F82575
R:Anonymous. The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: F82575
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-156 <S1M>
A:Cross-references: UNIPROT:Q9PB48; GB:AE004041; GB:AE003849; NID:g9107453; PIDN:AAF8509
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.B.; Kicajina, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.B.; Marques, M.V.; Martins, E. A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Tshunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XP2297

Query Match 2.8%; Score 7; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 VAALRAG 197
| | | | |
Db 136 VAALRAG 142

RESULT 13
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 6 - starfish (Asterina pectinifera)
C:Species: mitochondrion Asterina pectinifera
C>Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S70593
R:Asakawa, S.; Himeno, H.; Miura, K.; Watanabe, K.
Genetics 140, 1047-1060, 1995
A:Title: Nucleotide sequence and gene organization of the starfish Asterina pectinifera A:Reference number: S70589; MUID:95402698; PMID:7672576
A:Accession: S70593
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-162 <ASA>
A:Cross-references: UNIPROT:Q33817; EMBL:D16387; NID:9517153; PIDN:BAR03876.1; PID:g1066 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1993
C:Genetics:
A:Gene: mitochondrion
A:Genetic code: SGC8
A:Superfamily: NADH dehydrogenase (ubiquinone) chain 6
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 2.8%; Score 7; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 ILILVALV 164
| | | | |
Db 139 ILILVALV 145

RESULT 14
JQ0986
Lipid body-associated membrane protein - carrot
C:Species: Daucus carota (carrot)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: JQ0986
R:Hatzopoulos, P.; Franz, G.; Choy, L.; Sung, R.Z.
Plant Cell 2, 457-467, 1990
A:Title: Interaction of nuclear factors with upstream sequences of a lipid body membrane A:Reference number: JQ0986; MUID:93044488; PMID:2152171
A:Accession: JQ0986
A:Molecule type: DNA
A:Residues: 1-180 <HAT>
A:Cross-references: UNIPROT:Q43123; GB:S47635; NID:g259452; PIDN:AAB24078.1; PID:g259453 A:Experimental source: var. juwarot
C:Genetics:
A:Gene: DC59
A:Superfamily: oleosin
C:Keywords: membrane protein

Query Match 2.8%; Score 7; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GTLLFLA 23
| | | | |
Db 50 GTLLFLA 56

RESULT 15
AE0435
conserved hypothetical protein YP03580 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AE0435
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B. geno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AE0435
A:Molecule type: DNA
A>Status: preliminary
A:Residues: 1-181 <KUR>
A:Cross-references: UNIPROT:Q8ZB45; GB:AL590842; PIDN:CAC92809.1; PID:g15981502; GSPDB:G C:Genetics:
A:Gene: YP03580
C:Superfamily: conserved hypothetical protein NMB0355

Query Match 2.8%; Score 7; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 LAASLSA 28
| | | | |
Db 17 LAASLSA 23

Search completed: March 20, 2005, 10:18:52
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 09:50:19 ; Search time 92 Seconds
(without alignments)
1380.387 Million cell updates/sec

Title: US-10-080-522-1

Perfect score: 248

Sequence: 1 MQTCPLAPGHVSQALGTL.....PLGALELLSQPLFPYAADP 248

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	248	100.0	248	1	SCTM HUMAN
2	9	3.6	1709	2	Q8WV66 homo sapien
3	8	3.2	108	2	Q93H59 streptomyc
4	8	3.2	124	2	Q9YBBS aeropyrum p
5	8	3.2	159	2	Q8Y402 ophiura lut
6	8	3.2	163	2	Q8X1D5 uetillago ma
7	8	3.2	201	2	Q6AYS0 rattus norv
8	8	3.2	221	2	Q7NUR2 chromobacte
9	8	3.2	288	2	Q659M6 turnip mosa
10	8	3.2	304	2	Q64WT9 bacteroides
11	8	3.2	316	2	Q62X0 xenopus tro
12	8	3.2	332	2	Q7Y516 xanthomonas
13	8	3.2	384	2	Q83402 treponema p
14	8	3.2	521	2	Q897Y6 clostridium
15	8	3.2	545	2	Q82FX6 streptomyc
16	8	3.2	760	2	Q9N612 leishmania
17	8	3.2	2204	2	Q6Q356 newcastle d
18	8	3.2	2204	2	Q6RCH6 newcastle d
19	8	3.2	4912	2	Q94116 aureobasidi
20	7	2.8	39	2	Q9ERK6 mus musculu
21	7	2.8	51	2	Q9PD23 xylella fas
22	7	2.8	56	2	Q7Y191 oryza sativ
23	7	2.8	59	2	Q745H6 mycobacteri
24	7	2.8	68	1	TX1L OPINA
25	7	2.8	70	2	Q8UX21 opisthacant
26	7	2.8	70	2	Q8UX22 hepatitis c
27	7	2.8	70	2	Q8UX23 hepatitis c
28	7	2.8	70	2	Q8UX24 hepatitis c
29	7	2.8	70	2	Q8UX25 hepatitis c
30	7	2.8	70	2	Q8UX27 hepatitis c
31	7	2.8	70	2	Q8UX28 hepatitis c

32	7	2.8	70	2	Q8UX29	Q8ux29 hepatitis c
33	7	2.8	70	2	Q8UX30	Q8ux30 hepatitis c
34	7	2.8	70	2	Q8UX31	Q8ux31 hepatitis c
35	7	2.8	70	2	Q8UX32	Q8ux32 hepatitis c
36	7	2.8	70	2	Q8UX34	Q8ux34 hepatitis c
37	7	2.8	70	2	Q8UX35	Q8ux35 hepatitis c
38	7	2.8	70	2	Q8UX36	Q8ux36 hepatitis c
39	7	2.8	70	2	Q8UX37	Q8ux37 hepatitis c
40	7	2.8	70	2	Q8UX38	Q8ux38 hepatitis c
41	7	2.8	70	2	Q8UX39	Q8ux39 hepatitis c
42	7	2.8	70	2	Q8UX40	Q8ux40 hepatitis c
43	7	2.8	70	2	Q8UX41	Q8ux41 hepatitis c
44	7	2.8	70	2	Q8UX42	Q8ux42 hepatitis c
45	7	2.8	70	2	Q8UX43	Q8ux43 hepatitis c

ALIGNMENTS

RESULT 1

SCTM HUMAN STANDARD; PRT; 248 AA.
AC Q8WV66; Q00466;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Secreted and transmembrane protein 1 precursor (Protein K12).
GN Name=SECTM1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=98149980; PubMed=9480746; DOI=10.1006/geno.1997.5151;
RA Sents-Kesler K.A., Hale L.P., Kaufman R.E.;
RT "Identification and characterization of K12 (SECTM1), a novel human gene that encodes a Golgi-associated protein with transmembrane and secreted isoforms";
RL Genomics 47:327-340(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield F.S.N., Krzywinski M.I., Skalska U., Smal M.A., Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP INTERACTION WITH CD7.
RX MEDLINE=20119303; PubMed=10652336; DOI=10.1074/jbc.275.5.3431;
RA Lyman S.D., Escobar S., Rousseau A.-M., Armstrong A., Fanslow W.C.;
RT "Identification of CD7 as a cognate of the human K12 (SECTM1) protein";
RL J. Biol. Chem. 275:3431-3437(2000).
CC -!- SUBUNIT: Interacts with CD7.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable). Also

```

LT 2
59
Q93H59 PRELIMINARY; PRT; 1709 AA.
Q93H59;
01-DEC-2001 (TRENBLrel. 19, Created)
01-DEC-2001 (TRENBLrel. 19, Last sequence update)
25-OCT-2004 (TRENBLrel. 28, Last annotation update)
Non-ribosomal peptide synthetase (Putative non-ribosomal peptide
synthetase).
OrderedLocusNames=SAV3643;
Streptomyces avermitilis.
Bacteria: Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=33903;
[1] _____
SEQUENCE FROM N.A.
STRAIN=NA-4680;

```

```

RESULT 3
Q9YBB5
ID Q9YBB5 PRELIMINARY; PRT; 108 AA.
AC Q9YBB5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein APE1682.
GN OrderedLocustNames=APE1682;
OS Archaeum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN _SEQUENCE FROM N.A.
RP
RC STEAIN-K1;
RX MEDLINE=9310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,

```

RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix K1.",
 RL DNA Res. 6:83-101(1999)
 DR EMBL; AP000062; BAA80683.1; -.
 DR F01; F72549; F72549.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 108 AA; 11640 MW; 073BBD9F5EE95106 CRC64;

Query Match 3.2%; Score 8; DB 2; Length 108;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 TPRPLALV 224
 |||||
 DB 19 TPRPLALV 26

RESULT 4
 Q8R155 PRELIMINARY; PRT; 124 AA.
 AC Q8R155;
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE CDNA sequence BC025446.
 GN Name=BC025446;
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Liver;
 RX MEDLINE=2388257; PubMed=1477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Liver;
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC025446; AAB25446.1; -.
 DR MGI; MGI:2385015; BC025446.
 DR InterPro; IPR003632; LY-6 CD59.
 DR InterPro; IPR001526; LY6 UPAR.
 DR Pfam; PF00021; UPAR_LY6_1.
 DR ProDom; PD003128; LY-6 CD59; 1.
 DR SMART; SM00134; LU; 1.
 DR PROSITE; PS00983; LY6 UPAR; 1.
 SQ SEQUENCE 124 AA; 13401 MW; C53338BD584177B0 CRC64;

Query Match 3.2%; Score 8; DB 2; Length 124;
 Best Local Similarity 100.0%; Pred. No. 18;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 VFILLVAL 163
 |||||
 DB 11 VFILLVAL 18

RESULT 5
 Q6Y402 PRELIMINARY; PRT; 159 AA.
 AC Q6Y402;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE NADH dehydrogenase subunit 6.
 GN Name=nad6;
 OS Ophiura lutea.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
 OC Ophiuroidea; Ophiurida; Ophiurida; Ophiurina; Chilophiurina;
 OC Ophiuridae; Ophiura.
 OX NCBI_TaxID=224378;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15019608; DOI=10.1016/j.ympev.2003.07.005;
 RA Scouras A., Beckenbach K., Arndt A., Smith M.J.;
 RT "Complete mitochondrial genome DNA sequence for two ophiuroids and a
 RT holothuroid: the utility of protein gene sequence and gene maps in the
 RT analyses of deep deuterostome phylogeny.";
 RL Mol. Phylogenet. Evol. 31:50-65(2004).
 DR EMBL; AY184223; AAO65637.1; -.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u...; IEA.
 DR InterPro; IPR001457; Oxidored_q3.
 DR Pfam; PF00499; Oxidored_q3; 1.
 KW Mitochondrion.
 SQ SEQUENCE 159 AA; 17502 MW; 2758C60783FF5E86 CRC64;

Query Match 3.2%; Score 8; DB 2; Length 159;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 FILLVALV 164
 |||||
 DB 135 FILLVALV 142

RESULT 6
 Q8X1D5 PRELIMINARY; PRT; 163 AA.
 AC Q8X1D5;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Mig2-4.
 GN Name=mig2-4;
 OS Ustilago maydis (Smut fungus).
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
 OX NCBI_TaxID=5270;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21839640; PubMed=11849538;
 RA Basse C.W., Kolb S., Kahmann R.;
 RT "A maize-specifically expressed gene cluster in Ustilago maydis.";
 RL Mol. Microbiol. 43:75-93(2002).
 DR EMBL; AF429982; AAL67330.1; -.
 SQ SEQUENCE 163 AA; 17723 MW; 97A653FFA3E1E16 CRC64;

Query Match 3.2%; Score 8; DB 2; Length 163;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

NCEI_TaxID=536;
[...]
SEQUENCE FROM N.A.
STRAIN=ATCC 12472 / DSM 30191;
MEDLINE=22882880; PubMed=14500782; DOI=10.1073/pnas.1832124100;
RC Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.I.,
RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
RA Alves-Gomes J.A., Andrade E.M., Ataripi J.J., Araujo M.P.F.,
RA Astolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
RA Batista J.S., Belo A., van den Berg C., Bogio M., Bonatto S.,
RA Bordignon J., Brigo M.M., Brito C.A., Brocchi M., Burley H.A.,
RA Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
RA Carvalhal C.M.B., Cascado J.C.M., Cavada B.S., Chueire L.M.O.,
RA Creszynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
RA Fanicatti P., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
RA Ferri M.I.T., Franco G.R., Freitas N.S.A., Furian L.R.,
RA Gazinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
RA Gratapaglia D., Grisard E.C., Hanna E.S., Jardim S.N., Laurino J.,
RA Leoi L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
RA Madeira H.M.F., Manfio G.P., Maranhao A.Q., Martins W.S.,
RA di Mauro S.M.Z., de Medeiros S.R.B., Meisner R.V., Moreira M.A.M.,
RA Nascimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
RA Faixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
RA Pereira M., Pinto L.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
RA Ramalho-Neto C.E., Reis A.M.M., Rigto L.U., Rondinelli E.,
RA Santos E.B.P., Santos F.R., Schneider M.P.C., Seavanez H.N.,
RA Silva A.M.R., da Silva A.L.C., Silveira D.W., Souza E.M., Simoes I.C.,
RA Simon D., Soares C.M.A., Soares R.B.A., Souza K.R.L.,
RA Souza R.C., Steffens M.B.R., Steinzel M., Teixeira S.R., Urmenyi T.,
RA Vetore A., Wassem R., Zaha A., Simpson A.J.G.;
RT "The complete genome sequence of Chromobacterium violaceum reveals
remarkable and exploitable bacterial adaptability";
Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
-- SIMILARITY: Contains 1 response regulatory domain.
EMBL: AE016919; AAQ60305.1; --
DR HSPP; Q9WYN0; IKGS;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0000156; P:sensor perception; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p.p.); IEA.
DR InterPro; IPRO09059; bi resp regltr_C.
DR InterPro; IPRO01006; CheY like.
DR InterPro; IPRO01789; Response_reg.
DR InterPro; IPRO01867; Trans_reg_C.
DR Pfam; PF00072; Response_reg_1.
DR Pfam; PF00486; Trans reg C; 1.
DR ProDom; PD000039; Response_reg_1.
DR ProDom; PD000329; Trans reg C; 1.
DR PROSITE; PS00110; RESPONSE REGULATORY; 1.
KW Complete proteome; DNA-binding; Phosphorylation; Sensory transduction;
Transcription; Transcription regulation.
SQ SEQUENCE 221 AA; 24395 MW; 80A50R7536SBA30C CRC64;

Query Match 3.2%; Score 8; DB 2; Length 221;
Best Local Similarity 100.0%; Pred.No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 191 VAALRAGA 198
DB 88 VAALRAGA 95

RESULT 9
Q659M6 PRELIMINARY; PRT; 288 AA.

ID AC Q659M6 (TREMBlrel. 28, Created)
DT DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DT DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE DE Polyprotein (Fragment).
OS OS Turnip mosaic virus (strain Japanese) (TMUV).
OC OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.

OX NCBI_TaxID=12230;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NBXC2;
 RA Shi M., Zhou X.;
 RT "Genomic variation of turnip mosaic virus.";
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ831813; CAH42011.1; -;
 KW Coat protein; Polyprotein.
 FT NON_TER 1 1
 FT CHAIN <1 >288 coat protein.
 FT NON_TER 288 288
 SQ SEQUENCE 288 AA; 33104 MW; AA7364110079E314 CRC64;
 Query Match 3.2%; Score 8; DB 2; Length 288;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 188 QMKVAALR 195
 Db 240 QMKVAALR 247
 RESULT 10
 Q64WT9 PRELIMINARY; PRT; 304 AA.
 ID Q64WT9;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=BF1287;
 OS Bacteroides fragilis.
 OC Bacteria; Bacteroidetes.
 OC Bacteroidaceae; Bacteroides.
 OC Bacteroidales.
 OX NCBI_TaxID=817;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YCH46;
 RA Kuwahara T., Yamashita A., Hirakawa H., Nakayama H., Toh H., Okada N.,
 RA Kuwaha S., Hattori M., Hayashi T., Ohnishi Y.;
 RT "Genomic analysis of Bacteroides fragilis reveals extensive DNA
 RT inversions regulating cell surface adaptation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2004).
 DR EMBL; AP006841; BAD48037.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 304 AA; 33071 MW; 575A6D8FC7BEDAFA CRC64;
 Query Match 3.2%; Score 8; DB 2; Length 304;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 20 LFLAASLS 27
 Db 124 LFLAASLS 131
 RESULT 11
 Q6P2X0 PRELIMINARY; PRT; 316 AA.
 ID Q6P2X0
 AC Q6P2X0;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Hypothetical protein MGC76280.
 GN Name=MGC76280;
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8364;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Shat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Mullahy S.J.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallus J., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RL Klein S., Gerhard D.S.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RL Klein S., Gerhard D.S.;
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC064266; AAH64266.1; -;
 DR EMBL; BC080950; AAH80950.1; -;
 DR GO; GO:0004519; F:endonuclease activity; IEA.
 DR GO; GO:0006281; P:DNA repair; IEA.
 DR InterPro; IPR000097; ApEndonclisel.
 DR InterPro; IPR004808; ExoIII_xth.
 DR InterPro; IPR005135; Exo endo phos.
 DR Pfam; PF03372; Exo endo phos; 1.
 DR TIGRFAMs; TIGR00195; exoDNase_III; 1.
 DR TIGRFAMs; TIGR00633; xth; 1.
 DR PROSITE; PS00728; AP_NUCLEASE_F1_3; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 316 AA; 35955 MW; AB4E25B5215497B1 CRC64;
 Query Match 3.2%; Score 8; DB 2; Length 316;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 132 EVSGAEPQ 139
 Db 9 EVSGAEPQ 16
 RESULT 12
 Q7Y5I6 PRELIMINARY; PRT; 332 AA.
 ID Q7Y5I6
 AC Q7Y5I6;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE 31L.
 OS Xanthomonas oryzae bacteriophage Xp10.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales.
 OX NCBI_TaxID=232237;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22734642; PubMed=12850143; DOI=10.1016/S0022-2836(03)00634-X;
 RA Yuzenkova J., Nechaev S., Berlin J., Rogulja D., Kuznedelov K.,
 RA Inman R., Mushagian A., Severinov K.;
 RT "Genome of Xanthomonas oryzae bacteriophage Xp10: an odd T-odd
 RT phage.";

```

RL J. Mol. Biol. 330:735-748(2003).
RP [2]
RP SEQUENCE FROM N.A.
RP Nechaev S.Y., Yuzenkova Y.V., Berlin J., Regulja D.; Kurnedelov K.,
RA Inman R., Mushhegian A., Severinov K.V.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY298121; AAP58698.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003910; F:DNA ligase (ATP) activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR000977; DNA ligase.
DR InterPro; IPR008994; Nucleic acid OB.
DR Pfam; PF01068; DNA_ligase_A_M; 1.
SQ SEQUENCE 332 AA; 36551 MW; 6CB6C37F945D6CBF CRC64;

Query Match 3.2%; Score 8; DB 2; Length 332;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 LAASLSAQ 29
DB 193 LAASLSAQ 200
|||||
|1|1|1|1|1|

RESULT 13
OB3402 PRELIMINARY; PRT; 384 AA.
AC OB3402;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cell division protein (FtsW).
GN OrderedLocNames=FP0387;
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98332770; PubMed=9665876; DOI=10.1126/science.281.5375.375;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R.J., Gwinn M.L., Hickey E.K., Clayton R.A., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L., Peterson J.D.,
RA Khalak H.G., Richardson D.L., Howell J.K., Chidambaram M.,
RA Utterback T.R., McDonald L.A., Artach P., Bowman C., Cotton M.D.,
RA Fujii C., Garland S.A., Hatch B., Horst K., Roberts K.M., Sandusky M.,
RA Weidman J.F., Smith H.O., Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
spirochete."
RL Science 281:375-388(1998).
DR EMBL; AE001217; AAC65371.1; -.
DR PIR; E71331; E71331.
DR TIGR; TP0387; -.
DR GO; GO:0006021; C:integral to membrane; IEA.
DR GO; GO:0007049; F:cell cycle; IEA.
DR InterPro; IPR001182; Cell_cycle.
DR InterPro; IPR006162; Ppantne S.
DR Pfam; PF01098; FTSW RODA SPOVE; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 384 AA; 42714 MW; 01DF27E5E30B4F2E CRC64;

Query Match 3.2%; Score 8; DB 2; Length 384;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 VFILLVAL 163
DB 19 VFILLVAL 26
|||||
|1|1|1|1|1|

us-10-080-522-1.oli.rup

RESULT 14
OB37Y6 PRELIMINARY; PRT; 521 AA.
AC OB37Y6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative efflux pump component mtrF.
GN Name=mtrF; OrderedLocNames=CTC00584;
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Massachusetts / E88;
MEDLINE=22457253; PubMed=12552129; DOI=10.1073/pnas.0335853100;
RA Brueggemann H., Baeumer S., Fricke W.F., Wierzer A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
tetanus disease."
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
DR EMBL; AE015938; AAC035200.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004697; Put transpt_AbgT.
DR Pfam; PF03806; ABG_transport; 1.
KW Complete proteome.
SQ SEQUENCE 521 AA; 55829 MW; D1A9096671E58C8C CRC64;

Query Match 3.2%; Score 8; DB 2; Length 521;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 FILLVALV 164
DB 400 FILLVALV 407
|||||
|1|1|1|1|1|

RESULT 15
OB2FX6 PRELIMINARY; PRT; 545 AA.
AC OB2FX6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative membrane protein.
GN OrderedLocNames=SAV4126;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MA-4680;
MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Oonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
avermitilis: deducing the ability of producing secondary
metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=MA-4680;
MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial

```

RT microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005037; BAC71838.1;
DR GO: GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR004843; M-pesterase.
DR Pfam; PF00149; Metallophos; 1.
KW Complete proteome.
SQ SEQUENCE 545 AA; 57993 MW; BDCB0CA12396479 CRC64;

Query Match 3.2%; Score 8; DB 2; Length 545;
Best Local Similarity 100.0%; Pred.No. 67;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SPLGALEL 235
|||
Db 120 SPLGALEL 127

Search completed: March 20, 2005, 10:18:21
Job time : 96 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 06:49:53 ; Search time 85 Seconds
(without alignments)
1128.430 Million cell updates/sec

Title: US-10-080-522-1

Perfect score: 1308

Sequence: 1 MOTCPFLAPGHVSQALGTL.....PLGALELLSQPLFPFYAADP 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A. Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1308	100.0	248	2	Aaw48811 K12 prote
2	1308	100.0	248	2	Aay21846 Human sig
3	1308	100.0	248	4	Aab36658 Human K12
4	1308	100.0	248	8	Adj75366 Marker ge
5	1308	100.0	248	8	Adp24664 PRO polyp
6	916.5	70.1	183	6	Aao29897 Human org
7	720.5	55.1	149	7	Adb36336 Human imm
9	492	37.6	101	6	Aao29895 Human org
10	326	24.9	212	4	Aab36660 Mouse K12
11	326	24.9	212	8	Adj76167 Marker ge
12	325	24.8	212	8	Adj76168 Marker ge
13	294	22.5	107	6	Aao29898 Human org
14	286	21.9	55	3	Aab34671 Gene 35 h
15	279	21.3	55	3	Aab34672 Human sec
16	163	12.5	30	3	Aab34673 Gene 35 h
17	156	11.9	43	7	Adb36337 Human imm
18	100	7.6	874	4	Abb59013 Drosophil
19	96	7.3	435	6	Abr41273 Human DIT
20	96	7.3	442	4	Aab95569 Human pro
21	95.5	7.3	506	5	Adi16910 Cow NOVX
22	91	7.0	6620	7	Adj70485 Human hea
23	91	7.0	7968	6	Abg76187 Human ser
24	90.5	6.9	1171	7	Abo63781 Klebsiell
25	89	6.8	435	4	Aab61190 Mature hu

26	89	6.8	435	6	ABO32633	Ab032633 Secreted
27	89	6.8	435	7	ADB90720	Abd90720 Human INT
28	89	6.8	435	7	ADF71455	Adf71455 Human INT
29	89	6.8	435	8	ADQ10281	Adq10281 Human pol
30	89	6.8	455	4	AAAB61188	AAb61188 Human INT
31	89	6.8	455	6	ABO32631	Ab032631 Secreted
32	89	6.8	455	7	ADB90718	Abd90718 Human INT
33	89	6.8	455	7	ADF71453	Adf71453 Human INT
34	89	6.8	455	8	ADQ10279	Adq10279 Human pol
35	86	6.6	393	8	ADN95576	Adn95576 Novel hum
36	84.5	6.5	289	6	ABU15711	Abu15711 Protein e
37	84.5	6.5	324	2	AAR12427	Aar12427 Hybrid Pc
38	84.5	6.5	381	7	ABO79025	Ab079025 Pseudomon
39	84.5	6.5	455	8	ADQ59388	Adq59388 Human can
40	84.5	6.5	455	8	ADO76631	Ado76631 Human but
41	84.5	6.5	482	8	ADQ76633	Adq76633 Human but
42	84	6.4	444	8	ABM83615	Abm83615 Human dia
43	84	6.4	1161	8	ADQ67749	Adq67749 Novel hum
44	84	6.4	1482	6	ABR58656	AbR58656 Human can
45	84	6.4	1497	5	ABP69627	Abp69627 Human pol

ALIGNMENTS

RESULT 1
AAW48811
ID AAW48811 standard; protein; 248 AA.
XX AC AAW48811;
XX DT 26-OCT-1998 (first entry)
XX DE K12 protein.
XX KW DNA probe; CD7 Hs1 DNase hypersensitive site; mRNA northern blot;
KW human erythroleukemic; HEL; K562 cell line; clone; breast cancer;
KW ovarian cancer; malignant; tumour.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Misc-difference 239
FT FT /note= "encoded by CCAA"
XX PN WO9822502-A1.
XX PD 28-MAY-1998.
XX PF 24-NOV-1997; 97WO-US021517.
XX PR 22-NOV-1996; 96US-00755559.
XX PA (UYDU-) UNIV DUKE.
XX KA Kaufman RE, Slentz-Kesler KA;
XX WPI; 1998-312415/27.
XX N-PSDB; AAV32446.
XX PT New isolated K12 protein gene - which is over expressed in certain
XX neoplastic cells, used to develop products for tumour detection and
XX treatment.
XX PS Claim 3; Fig 1; 44pp; English.
XX CC This present sequence represents the K12 protein, the gene for which has
XX been located on chromosome 17q25. To obtain this genes cDNA sequence a
XX 500 bp DNA probe, which can be located just upstream of the CD7 Hs1 DNase
XX hypersensitive site, was used against a mRNA northern blot. From this a
XX 1.8kb transcript was detected in the human erythroleukemic cell line HEL.
XX The probe was then used to screen a human erythroleukemic cell line K562
XX cDNA library, from which several clones were identified and isolated that

CC constituted a 1.9kb cDNA. This cDNA was designated K12 and was found to
CC have a single open reading frame as well as being in the same orientation
CC as CD7. The K12 gene was found to be expressed in both breast and ovarian
CC cancer cells at a much higher level than any other malignant or normal
CC tissue that was examined, thus enabling the K12 to be a useful protein in
CC tumour detection and treatment
XX
SQ Sequence 248 AA;
Query Match 100.0%; Score 1308; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.8e-133;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 M QTCPLAFPGHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSWGENTVMSCNISNA 60
DB 1 M QTCPLAFPGHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSWGENTVMSCNISNA 60
QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
QY 121 GHQRNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVNFAYWRCRCQORREK 180
DB 121 GHQRNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVNFAYWRCRCQORREK 180
QY 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSETPRPLALVFKPSPGLALELLSPQ 240
DB 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSETPRPLALVFKPSPGLALELLSPQ 240
QY 241 LFPYAADP 248
DB 241 LFPYAADP 248
RESULT 3
AAB36658
ID AAB36658 standard; protein; 248 AA.
XX
AC AAB36658;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human K12 protein sequence SEQ ID NO:4.
XX
KW Human; CD7; K12; cognate ligand; cluster of differentiation; cancer;
KW identification; inhibiting T cell proliferation; HIV; infection;
KW activating natural killer cell proliferation; leukaemia; lymphoma;
KW sepsis; graft versus host disease; autoimmune disease; arthritis;
KW multiple sclerosis; rheumatoid arthritis; psoriatic arthritis; lupus;
KW scleroderma; psoriasis; atopic dermatitis; type 1 diabetes mellitus;
KW Hashimoto's thyroiditis; pernicious anaemia; Addison's disease; uveitis;
KW myasthenia gravis; psoriasis; Guillain-Barre Syndrome; Grave's disease;

PT Human signal-peptide containing protein coding sequences used to treat
PT cancer and immune responses.
XX
PS Claim 1; Page 79-80; 99p; English.
XX
CC The invention provides human signal-peptide containing proteins (SIGP)
CC (AAV21841-855) and polynucleotides (AAK82076-90) encoding the proteins. A
CC host cell containing a vector comprising SIGP DNA can be used to produce
CC the SIGP protein. The SIGP protein can be used, in conjunction with a
CC pharmaceutical carrier to treat or prevent a cancer. An antagonist of the
CC SIGP protein can be used to treat or prevent a cancer or an immune
CC response. The cancers that can be treated or prevented include sarcomas,
CC adenocarcinomas, leukemia's, lymphomas, melanomas, teratocarcinomas,
CC myelomas and cancers of the adrenal gland, bladder, bone, bone marrow,
CC brain, breast, cervix, gall bladder, ganglia, pancreas, parathyroid, penis,
CC heart, kidney, liver, lung, muscle, ovary, spleen, testis, thymus, thyroid, and
CC prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and
CC uterus. The immune responses that can be treated or prevented include,
CC AIDS, Addison's disease, adult respiratory distress syndrome, allergies,
CC anemia, asthma, atherosclerosis, bronchitis, cholecystitis, Crohn's
CC disease, ulcerative colitis, atopic dermatitis, dermatomyositis, diabetes
CC mellitus, emphysema, atrophic gastritis, glomerulonephritis, Grave's
CC disease, gout, hyperesoinophilia, irritable bowel syndrome, lupus
CC erythematosus, multiple sclerosis, myasthenia gravis, myocardial or
CC pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,
CC polymyositis, rheumatoid arthritis, scleroderma, Sjogren's syndrome, and
CC autoimmune thyroiditis, complications of cancer, infections, and trauma
XX
SQ Sequence 248 AA;
Query Match 100.0%; Score 1308; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.8e-133;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 M QTCPLAFPGHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSWGENTVMSCNISNA 60
DB 1 M QTCPLAFPGHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSWGENTVMSCNISNA 60
QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
DB 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 120
QY 121 GHQRNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVNFAYWRCRCQORREK 180
DB 121 GHQRNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVNFAYWRCRCQORREK 180
QY 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSETPRPLALVFKPSPGLALELLSPQ 240
DB 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSETPRPLALVFKPSPGLALELLSPQ 240
QY 241 LFPYAADP 248
DB 241 LFPYAADP 248
RESULT 2
AAV21846
ID AAV21846 standard; protein; 248 AA.
XX
AC AAV21846;
XX
DT 20-SEP-1999 (first entry)
XX
DE Human signal peptide-containing protein (SIGP) (clone ID 1747327).
XX
KW Signal-peptide containing protein; SIGP; human; cancer; immune response;
KW adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma; AIDS;
KW Addison's disease; adult respiratory distress syndrome; allergy; anemia;
KW asthma; atherosclerosis; bronchitis; cholecystitis; Crohn's disease;
KW ulcerative colitis; atopic dermatitis; dermatomyositis; emphysema;
KW diabetes mellitus; atrophic gastritis; glomerulonephritis; gout; trauma;
KW Grave's Disease; hyperesoinophilia; irritable bowel syndrome; infection;
KW lupus erythematosus; multiple sclerosis; myasthenia gravis; inflammation;
KW osteoarthritis; osteoporosis; pancreatitis; polymyositis; scleroderma;
KW rheumatoid arthritis; Sjogren's syndrome; autoimmune thyroiditis.
XX
OS Homo sapiens.
XX
FN WO9933981-A2.
XX
PD 08-JUL-1999.
XX
PF 22-DEC-1998; 98WO-US027598.
XX
PR 31-DEC-1997; 97US-00002485.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Lal P, Hillman JL, Corley NC, Guegler KJ, Baughn MR, Sather SK;
PI Shah P;
XX
XX WPI; 1999-430242/36.
DR
DR N-PSDB; AAK82081.
XX

KW systemic lupus erythematosus; dermatomyositis; asthma; eczema;
KW atopic dermatitis; contact dermatitis; eczematous dermatitis;
KW seborrheic dermatitis; rhinitis.

XX Homo sapiens.

XX WO200073333-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US014612.

XX 28-MAY-1999; 99US-0136450P.

XX (IMMV) IMMUNEX CORP.

XX Lyman SD, Fanslow WC;

XX WPI; 2001-061511/07.

XX N-PSDB; AAC88152.

XX Stimulating intracellular signaling of CD7 comprises contacting a cell
PT expressing CD7 with recombinant K12 protein, the cognate ligand of CD7,
PT to inhibit T cell proliferation and/or activate natural killer cell
PT proliferation.

XX Claim 2; Page 38-39; 42pp; English.

XX The present invention describes a method for stimulating (S) the
CC intracellular signaling of CD (Cluster of differentiation) 7 comprising
CC contacting a cell that expresses CD7 with a recombinant K12 protein (I),
CC the cognate ligand of CD7. (S) is useful for inhibiting T cell
CC proliferation and/or activating NK (natural killer) cell proliferation
CC and/or inducing NK toxicity in a mammal which involves administration of
CC K12 protein. It is also used for treating HIV-1 infection, cancer (T cell
CC leukaemia, acute lymphocytic leukaemia, cutaneous T cell lymphoma),
CC bacterial and viral infections, mediated by CD7. In the case of treating
CC T cell leukaemia the soluble K12 protein is covalently attached to a
CC toxin. A disease mediated by CD7 such as sepsis, graft versus host
CC disease due to transplantation, autoimmune diseases, multiple sclerosis,
CC arthritis, rheumatoid arthritis, psoriatic arthritis, scleroderma, lupus,
CC psoriasis, atopic dermatitis, type 1 diabetes mellitus, Hashimoto's
CC thyroiditis, pernicious anaemia, Addison's disease, myasthenia gravis,
CC uveitis, psoriasis, Guillain-Barre Syndrome, Grave's disease, systemic
CC lupus erythematosus and dermatomyositis, asthma, eczema, atopic
CC dermatitis, contact dermatitis, other eczematous dermatides, seborrheic
CC dermatitis, and rhinitis is also treated by administering a K12
CC antagonist (neutralising antibody). The present sequence represents the
CC human K12 protein, which is given in the exemplification of the present
XX invention

XX Sequence 248 AA;

Query Match 100.0%; Score 1308; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.8e-133;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCLAPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

DB 1 MOTCLAPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

QY 61 FSHVNIKLRAHQESAIENEVAPGYFSRDMQLOVQGGVAQLVIKGDARDSHAGLYMHLV 120

DB 61 FSHVNIKLRAHQESAIENEVAPGYFSRDMQLOVQGGVAQLVIKGDARDSHAGLYMHLV 120

QY 121 GQRNRRQVTLVSGAEPQSPADPTGFWPVPVAVTAVFILLVALVNFAYWRCRCQORREK 180

DB 121 GQRNRRQVTLVSGAEPQSPADPTGFWPVPVAVTAVFILLVALVNFAYWRCRCQORREK 180

QY 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPGLGALLSPQ 240

DB 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPDSEPTPRPLALVFKPSPGLGALLSPQ 240

QY 241 LFPYAADP 248

|||||

241 LFPYAADP 248

RESULT 4

ADJ75366

ID ADJ75366 standard; protein; 248 AA.

XX AC ADJ75366;

XX 20-MAY-2004 (first entry)

XX Marker gene related amino acid sequence SEQ ID NO:618.

XX bronchial asthma; chronic obstructive pulmonary disease;

XX respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker.

XX OS Homo sapiens.

XX EPI394274-A2.

XX 03-MAR-2004.

XX 04-AUG-2003; 2003EP-00254857.

XX 06-AUG-2002; 2002JP-00229312.

XX 20-MAR-2003; 2003JP-00077212.

XX (GENO-) GENOX RES INC.

XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuohara K;

XX WPI; 2004-193155/19.

XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.

XX Example 11; SEQ ID NO 618; 241pp; English.

XX The present invention describes a method of testing for bronchial asthma
or chronic obstructive pulmonary disease. The method comprises
determining the expression level of a marker gene in a biological sample
from a subject, comparing the expression level determined with the
expression level of the marker gene in a biological sample from a healthy
subject, and judging whether the subject has bronchial asthma or chronic
obstructive pulmonary disease. The marker gene comprises: (a) a group of
genes (S1) whose expression levels increase when respiratory epithelial
cells are stimulated with interleukin-13; or (b) a group of genes (S2)
whose expression levels decrease when respiratory epithelial cells are
stimulated with interleukin-13. Also described: (1) a reagent (I) for
testing for bronchial asthma or chronic obstructive pulmonary disease;
(2) a kit for screening for a candidate compound for a therapeutic agent
to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
an animal model for bronchial asthma or chronic obstructive pulmonary
disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
method for producing an animal model for bronchial asthma or chronic
obstructive pulmonary disease; (6) a therapeutic agent for bronchial
asthma or chronic obstructive pulmonary disease, comprising the compound,
a marker gene or an antisense nucleic acid corresponding to a portion of
the marker gene, a ribozyme, a polynucleotide that suppresses the
expression of the gene through an RNAi effect or an antibody recognising
a protein encoded by a marker gene; and (7) a DNA chip for testing for
bronchial asthma or a chronic obstructive pulmonary disease, on which a
probe has been immobilised to assay a marker gene. (I) has respiratory
and antiasthmatic activities, and can be used in gene therapy. The method
is useful for testing for or screening for a therapeutic agent for
bronchial asthma or chronic obstructive pulmonary disease. The present
sequence is used in the exemplification of the present invention.

XX

SQ Sequence 248 AA;

Query Match 100.0%; Score 1308; DB 8; Length 248;

Best Local Similarity 100.0%; Pred. No. 1.8e-133;

Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCTPLAPFGHVSQALGTLTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

Db 1 MOTCTPLAPFGHVSQALGTLTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGQVLQVGGVAQLVKGARDSHAGLYMWHLV 120

Db 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGQVLQVGGVAQLVKGARDSHAGLYMWHLV 120

QY 121 GHQRNNQVTLVSGAEPQAPDTGFWPVPVAVTAVFILLVALVFWAYRCRCQORREK 180

Db 121 GHQRNNQVTLVSGAEPQAPDTGFWPVPVAVTAVFILLVALVFWAYRCRCQORREK 180

QY 181 KFFLLEPQMKVAALRAGAQQGLSRASAEI WTPDSEPTPRPLALVFKPSPLGALLELLSPQ 240

Db 181 KFFLLEPQMKVAALRAGAQQGLSRASAEI WTPDSEPTPRPLALVFKPSPLGALLELLSPQ 240

QY 241 LFPYAADP 248

Db 241 LFPYAADP 248

RESULT 5

ADP24664

ID ADP24664 standard; protein; 248 AA.

XX

AC ADP24664;

DT 18-NOV-2004 (first entry)

XX

DE PRO polypeptide SEQ ID NO:1842.

XX

KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;

KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;

KW antiaesthetic; hepatotropic; respiratory; gene therapy; immune system.

XX

OS Unidentified.

XX

FN WO2004041170-A2.

XX

XX 21-MAY-2004.

XX

PF 30-OCT-2003; 2003WO-US034312.

XX

PR 01-NOV-2002; 2002US-0423394P.

XX

XX (GETH) GENENTECH INC.

XX

PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;

PI Wu TD;

XX

DR WPI; 2004-419628/39.

DR N-PSDB; ADP24663.

XX

XX New PRO polypeptides and polynucleotides, useful for treating e.g.

PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated

PT renal disease, or demyelinating diseases of the central or peripheral

PT nervous system.

XX

PS Claim 7; SEQ ID NO 1842; 2940pp; English.

XX

CC The invention relates to a novel isolated nucleic acid and the PRO

CC polypeptide encoded by it. A protein of the invention has

CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,

CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,

CC antiaesthetic, hepatotropic, and respiratory activity. A polynucleotide

CC of the invention may have a use in gene therapy. The PRO polypeptide, its

CC agonist, antagonist, or antibody that specifically binds to the

CC

polypeptide is useful for treating an immune related disorder such as

systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,

juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an

idiopathic inflammatory myopathy, Sjogren's syndrome, autoimmune

vasculitis, sarcoidosis, autoimmune haemolytic anaemia, systemic

thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal

disease, a demyelinating disease of the central or peripheral nervous

system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,

a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary

disease, infectious or autoimmune chronic active hepatitis, primary

biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,

inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's

disease, an autoimmune or immune-mediated skin disease, a bullous skin

disease, erythema multiforme, contact dermatitis, psoriasis, an allergic

disease, asthma, allergic rhinitis, atopic dermatitis, food

hypersensitivity, urticaria, an immunologic disease of the lung,

eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity

pneumonitis, a transplantation associated disease, graft rejection or

graft-versus-host disease. The present sequence represents a PRO protein

of the invention.

XX

XX Sequence 248 AA;

QY

Query Match 100.0%; Score 1308; DB 8; Length 248;

Best Local Similarity 100.0%; Pred. No. 1.8e-133;

Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTCTPLAPFGHVSQALGTLTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

Db 1 MOTCTPLAPFGHVSQALGTLTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGQVLQVGGVAQLVKGARDSHAGLYMWHLV 120

Db 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGQVLQVGGVAQLVKGARDSHAGLYMWHLV 120

QY 121 GHQRNNQVTLVSGAEPQAPDTGFWPVPVAVTAVFILLVALVFWAYRCRCQORREK 180

Db 121 GHQRNNQVTLVSGAEPQAPDTGFWPVPVAVTAVFILLVALVFWAYRCRCQORREK 180

QY 181 KFFLLEPQMKVAALRAGAQQGLSRASAEI WTPDSEPTPRPLALVFKPSPLGALLELLSPQ 240

Db 181 KFFLLEPQMKVAALRAGAQQGLSRASAEI WTPDSEPTPRPLALVFKPSPLGALLELLSPQ 240

QY 241 LFPYAADP 248

Db 241 LFPYAADP 248

RESULT 6

AAO29897

ID AAO29897 standard; protein; 183 AA.

XX

AC AAO29897;

XX

DT 03-SEP-2003 (first entry)

XX

DE Human organelle-associated protein (ORGA)-10.

XX

KW Human; organelle-associated protein; ORGA; cell proliferative disorder;

KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;

KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;

KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;

KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;

KW renal disorder; ovulatory defect; teratogenesis; pick's disease; cancer;

KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;

KW indigestion; gastritis; hypertension; gene therapy; prostatic; anorexia;

KW infection; urethritis; prostatitis.

XX

OS Homo sapiens.

XX

XX WO2003044171-A2.

XX

PD 30-MAY-2003.


```

XX PF 15-NOV-2002; 2002WO-US036807.
XX PR 16-NOV-2001; 2001US-0332384P.
XX PR 13-DEC-2001; 2001US-0341187P.
XX PR 23-JAN-2002; 2002US-0351151P.
XX PR 27-FEB-2002; 2002US-0360269P.
XX PR 05-APR-2002; 2002US-0370637P.
XX PR 14-JUN-2002; 2002US-0388946P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX
XX Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
PI Richardson TW, Gorvad AB, Sprague WW, Elliott VS, Khare R;
PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
PI Hafalia AJA, Bulloch S;
XX
XX WPI: 2003-457603/43.
XX DR N-PSDB; AAL60550.
XX
XX New organelle-associated proteins and polynucleotides, useful for
PT diagnosing, treating and/or preventing cell proliferative, reproductive,
PT gastrointestinal, neurological, urologic, and renal disorders.
XX
XX Claim 1; Page 166-167; 194pp; English.
XX
XX The invention relates to human organelle-associated proteins (ORGA) and
CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
CC for diagnosing, treating and/or preventing cell proliferative disorders
CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
CC disease, Pick's disease, Huntington's disease or dementia), urologic
CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
CC amyloidosis, renal failure, Addison's disease or hypertension). The
CC invention is also useful in gene therapy. The present sequence is human
CC ORGA protein
XX
XX Sequence 183 AA;
SQ
Query Match 70.1%; Score 916.5; DB 6; Length 183;
Best Local Similarity 73.4%; Pred. No. 4.9e-91;
Matches 182; Conservative 0; Mismatches 1; Indels 65; Gaps 1;
QY 1 MOTCLAPFCHVSQALGTLFLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNINA 60
DB 1 MOTCLAPFCHVSQALGTLFLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNINA 60
QY 61 FSHVNIKLRHQGSAIFNEVAPGVFSRDGWLQVGGVAQLVIKGDASHAGLYMHLV 120
DB 61 FSHVNIKLRHQGSAIFNEVAPGVFSRDGWLQVGGVAQLVIKGDASHAGLYMHLV 120
QY 121 GHQRNRQVTLVSGAEPQSDPTDGTGFWPVPVAVTAVTLLVALVNFVWRCRCQORREK 180
DB 121 GH----- 122
QY 181 KFFLEPQKVAALRAGAQOGLSRASAEIWTDPDSFTPRPLALVPKPSPLGALLELLSQP 240
DB 123 -----QTKVAALRAGAQOGLSRASAEIWTDPDSFTPRPLALVPKPSPLGALLELLSQP 175
QY 241 LFPYAADP 248
DB 176 LFPYAADP 183

```

RESULT 7
ADB36336

ID ADB36336 standard; protein; 149 AA.

XX ADB36336;
AC ADB36336;

```

XX DT 04-DEC-2003 (first entry)
XX
XX Human immune response associated protein IRAP-7 SEQ ID NO:7.
XX
XX human; immune response associated protein; IRAP; anti-HIV; anti-allergic;
XX antianemic; antiasthmatic; antiarteriosclerotic; dermatological;
XX antiinflammatory; antidiabetic; nephrotropic; antithyroid; thyromimetic;
XX immunosuppressive; antirheumatic; antiarthritic; osteopathic;
XX antibacterial; virucide; antiparasitic; protozoacide; fungicide;
XX cerebroprotective; neuroprotective; nootropic; antiparkinsonian;
XX antipsoriatic; cytostatic; cardiant; gene therapy;
XX immune system disorder; neurological disorder; developmental disorder;
XX muscle disorder; cell proliferative disorder; AIDS; allergy; anaemia;
XX asthma; atherosclerosis; contact dermatitis; Crohn's disease; diabetes;
XX glomerulonephritis; Grave's disease; Hashimoto's thyroiditis;
XX multiple sclerosis; rheumatoid arthritis; osteoporosis;
XX systemic lupus erythematosus; infection; stroke; Alzheimer's disease;
XX Parkinson's disease; psoriasis; cancer; cardiomyopathy.
XX
XX Homo sapiens.
XX
XX WO2003074726-A2.
XX
XX 12-SEP-2003.
XX
XX 28-FEB-2003; 2003WO-US006307.
XX
XX 01-MAR-2002; 2002US-0361088P.
XX 27-MAR-2002; 2002US-0368494P.
XX 10-MAY-2002; 2002US-0379876P.
XX 28-JUN-2002; 2002US-0392641P.
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Chawla NK, Becha SD, Jin P, Tran UK, Kable AE, Lee S;
XX Marquis JP, Emerling BM, Ison CH, Ramkumar J, Chang H;
XX WPI: 2003-722079/68.
XX DR N-PSDB; ADB36352.
XX
XX New human immune response associated proteins and polynucleotides, useful
PT for diagnosing, preventing or treating diseases or conditions associated
PT with aberrant enzyme expression, e.g. cancer, AIDS, atherosclerosis or
PT stroke.
XX
XX Claim 1; Page 143; 158pp; English.
XX
XX ADB36346 to ADB36361 encode the human immune response associated proteins
XX given in ADB36330 to ADB36345, designated IRAP-1 to IRAP-16. IRAP
XX sequences have anti-HIV, anti-allergic, antianemic, antiasthmatic,
XX antiarteriosclerotic, dermatological, antiinflammatory, antidiabetic,
XX nephrotropic, antithyroid, thyromimetic, immunosuppressive, virucide,
XX antirheumatic, antiarthritic, osteopathic, antibacterial, virucide,
XX antiparasitic, protozoacide, fungicide, cerebroprotective,
XX neuroprotective, nootropic, antiparkinsonian, antipsoriatic, cytostatic
XX and cardiant activities, and can be used in gene therapy. The IRAP
XX proteins and polynucleotides can be used in diagnosing, preventing or
XX treating diseases or conditions associated with the decreased expression
XX or overexpression of IRAP, such as immune system, neurological,
XX developmental, muscle or cell proliferative disorders. The disorders may
XX include AIDS, allergies, anaemia, asthma, atherosclerosis, contact
XX dermatitis, Crohn's disease, diabetes, glomerulonephritis, Grave's
XX disease, Hashimoto's thyroiditis, multiple sclerosis, rheumatoid
XX arthritis, osteoporosis, systemic lupus erythematosus, infections (e.g.
XX bacterial, viral, fungal, parasitic or protozoal), stroke, Alzheimer's
XX disease, Parkinson's disease, psoriasis, cancer and cardiomyopathy. They
XX are also useful in assessing the effects of exogenous compounds on the
XX expression of nucleic acid and amino acid sequences of IRAP. The IRAP or
XX its fragments are useful in screening compounds for effectiveness as
XX agonist or antagonist of the polypeptides, or in altering the expression
XX of the target polynucleotide and compounds that specifically bind to or
XX modulate the activity of the polypeptide.

```

CC for diagnosing, treating and/or preventing cell proliferative disorders
 CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
 CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
 CC ovulatory defects, ectopic pregnancy or teratogenesis), gastritis, indigestion,
 CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
 CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
 CC disease, Pick's disease, Huntington's disease or dementia), urologic
 CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
 CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
 CC amyloidosis, renal failure, Addison's disease or hypertension). The
 CC invention is also useful in gene therapy. The present sequence is human
 CC ORGA protein
 CC
 CC SQ Sequence 149 AA;
 SQ
 Query Match 55.1%; Score 720.5; DB 7; Length 149;
 Best Local Similarity 97.2%; Pred. No. 8e-70; Indels 1; Gaps 1;
 Matches 138; Conservative 1; Mismatches 2;
 QY 1 MOTCPPLAFPGHVSQALGTLFLAASLSAQNGWDSPTCTEGVSVSWGENTVMSCNISNA 60
 DB 1 MOTCPPLAFPGHVSQALGTLFLAASLSAQNGWDSPTCTEGVSVSWGENTVMSCNISNA 60
 QY 61 FSHVNIKLRAHQESAI FNEVAPGFYFRDGMQLQVGGVAQLVKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGFYFRDGMQLQVGGVAQLVKGARDSHAGLYMHLV 120
 QY 121 GHORNNRQVTLVSGAE-PSA 141
 DB 121 GHORNNRQVTLVSGAAVPSNA 142
 RESULT 8
 AAO29896
 ID AAO29896 standard; protein; 162 AA.
 XX
 AC AAO29896;
 XX
 DT 03-SEP-2003 (first entry)
 XX
 DE Human organelle-associated protein (ORGA) -9.
 XX
 KW Human; organelle-associated protein; ORGA; cell proliferative disorder;
 KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
 KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
 KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
 KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
 KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
 KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
 KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
 KW infection; urethritis; prostatitis.
 XX
 OS Homo sapiens.
 XX
 FN WO2003044171-A2.
 XX
 PD 30-MAY-2003.
 XX
 PF 15-NOV-2002; 2002WO-US036807.
 XX
 PR 16-NOV-2001; 2001US-0332384P.
 PR 13-DEC-2001; 2001US-0341187P.
 PR 23-JAN-2002; 2002US-0351151P.
 PR 27-FEB-2002; 2002US-0360269P.
 PR 05-APR-2002; 2002US-0370637P.
 PR 14-JUN-2002; 2002US-0388946P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
 PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
 PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
 PI Hafalia AJA, Bulloch S;
 XX
 DR WPI; 2003-457603/43.
 DR N-P8DB; AAL60549.
 XX
 PT New organelle-associated proteins and polynucleotides, useful for
 PT diagnosing, treating and/or preventing cell proliferative, reproductive,
 PT gastrointestinal, neurological, urologic, and renal disorders.
 XX
 PS Claim 1; Page 166; 194pp; English.
 XX
 CC The invention relates to human organelle-associated proteins (ORGA) and
 CC nucleic acid molecules encoding such proteins. ORGA sequences are useful

CC for diagnosing, treating and/or preventing cell proliferative disorders
 CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
 CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
 CC ovulatory defects, ectopic pregnancy or teratogenesis), gastritis, indigestion,
 CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
 CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
 CC disease, Pick's disease, Huntington's disease or dementia), urologic
 CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
 CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
 CC amyloidosis, renal failure, Addison's disease or hypertension). The
 CC invention is also useful in gene therapy. The present sequence is human
 CC ORGA protein
 CC
 CC SQ Sequence 162 AA;
 SQ
 Query Match 55.0%; Score 719.5; DB 6; Length 162;
 Best Local Similarity 89.1%; Pred. No. 1.2e-69; Indels 11; Gaps 1;
 Matches 139; Conservative 0; Mismatches 6;
 QY 1 MOTCPPLAFPGHVSQALGTLFLAASLSAQNGWDSPTCTEGVSVSWGENTVMSCNISNA 60
 DB 1 MOTCPPLAFPGHVSQALGTLFLAASLSAQNGWDSPTCTEGVSVSWGENTVMSCNISNA 60
 QY 61 FSHVNIKLRAHQESAI FNEVAPGFYFRDGMQLQVGGVAQLVKGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAI FNEVAPGFYFRDGMQLQVGGVAQLVKGARDSHAGLYMHLV 120
 QY 121 GHORNNRQVTLVSGAE-PSA 141
 DB 121 GHORNNRQVTLVSGAAVPSNA 142
 RESULT 9
 AAO29895
 ID AAO29895 standard; protein; 101 AA.
 XX
 AC AAO29895;
 XX
 DT 03-SEP-2003 (first entry)
 XX
 DE Human organelle-associated protein (ORGA) -8.
 XX
 KW Human; organelle-associated protein; ORGA; cell proliferative disorder;
 KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
 KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
 KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
 KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
 KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
 KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
 KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
 KW infection; urethritis; prostatitis.
 XX
 OS Homo sapiens.
 XX
 FN WO2003044171-A2.
 XX
 PD 30-MAY-2003.
 XX
 PF 15-NOV-2002; 2002WO-US036807.
 XX
 PR 16-NOV-2001; 2001US-0332384P.
 PR 13-DEC-2001; 2001US-0341187P.
 PR 23-JAN-2002; 2002US-0351151P.
 PR 27-FEB-2002; 2002US-0360269P.
 PR 05-APR-2002; 2002US-0370637P.
 PR 14-JUN-2002; 2002US-0388946P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;
 PI Richardson TW, Gorvad AE, Sprague WW, Elliott VS, Khare R;
 PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;
 PI Hafalia AJA, Bulloch S;
 XX
 DR WPI; 2003-457603/43.
 DR N-P8DB; AAL60549.
 XX
 PT New organelle-associated proteins and polynucleotides, useful for
 PT diagnosing, treating and/or preventing cell proliferative, reproductive,
 PT gastrointestinal, neurological, urologic, and renal disorders.
 XX
 PS Claim 1; Page 166; 194pp; English.
 XX
 CC The invention relates to human organelle-associated proteins (ORGA) and
 CC nucleic acid molecules encoding such proteins. ORGA sequences are useful

```

XX DR WPI; 2003-457603/43.
XX DR N-PSDB; AAL60548.
XX PT New organelle-associated proteins and polynucleotides, useful for
XX PT diagnosing, treating and/or preventing cell proliferative, reproductive,
XX PT gastrointestinal, neurological, urologic, and renal disorders.
XX PS Claim 1; Page 165-166; 194pp; English.
XX CC The invention relates to human organelle-associated proteins (ORGA) and
XX CC nucleic acid molecules encoding such proteins. ORGA sequences are useful
XX CC for diagnosing, treating and/or preventing cell proliferative disorders
XX CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers
XX CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,
XX CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal
XX CC disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion,
XX CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's
XX CC disease, Pick's disease, Huntington's disease or dementia), urologic
XX CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper
XX CC tract infections) and renal disorders (e.g. glomerulonephritis, renal
XX CC amyloidosis, renal failure, Addison's disease or hypertension). The
XX CC invention is also useful in gene therapy. The present sequence is human
XX CC ORGA protein
XX SQ Sequence 101 AA;
    Query Match 37.6%; Score 492; DB 6; Length 101;
    Best Local Similarity 100.0%; Pred. No. 3.5e-45;
    Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSWGENTVMSCNISNA 60
Db 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSWGENTVMSCNISNA 60
QY 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGMQ 92
Db 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGMQ 92
RESULT 10
AAB36660
ID AAB36660 standard; protein; 212 AA.
XX AC AAB36660;
XX DT 13-MAR-2001 (first entry)
XX DE Mouse K12 protein sequence SEQ ID NO:8.
XX KW Mouse; CD7; K12; cognate ligand; cluster of differentiation; cancer;
XX KW identification; inhibiting T cell proliferation; HIV; infection;
XX KW activating natural killer cell proliferation; leukaemia; lymphoma;
XX KW sepsis; graft versus host disease; autoimmune disease; arthritis;
XX KW multiple sclerosis; rheumatoid arthritis; psoriatic arthritis; lupus;
XX KW scleroderma; psoriasis; atopic dermatitis; type I diabetes mellitus;
XX KW Hashimoto's thyroiditis; pernicious anaemia; Addison's disease; uveitis;
XX KW myasthenia gravis; psoriasis; Guillain-Barre Syndrome; Grave's disease;
XX KW systemic lupus erythematosus; dermatomyositis; asthma; eczema;
XX KW atopic dermatitis; contact dermatitis; eczematous dermatitis;
XX KW seborrheic dermatitis; rhinitis.
XX OS Mus sp.
XX PN WO200073333-A2.
XX PD 07-DEC-2000.
XX PF 26-MAY-2000; 2000WO-US014612.
XX PR 28-MAY-1999; 99US-0136450P.
XX PA (IMMUNEX CORP.
XX PT Stimulating intracellular signaling of CD7 comprises contacting a cell
XX PT expressing CD7 with recombinant K12 protein, the cognate ligand of CD7,
XX PT to inhibit T cell proliferation and/or activate natural killer cell
XX PT proliferation.
XX PS Claim 32; Page 42; 42pp; English.
XX CC The present invention describes a method for stimulating (S) the
XX CC intracellular signalling of CD (cluster of differentiation) 7 comprising
XX CC contacting a cell that expresses CD7 with a recombinant K12 protein (1),
XX CC the cognate ligand of CD7. (S) is useful for inhibiting T cell
XX CC proliferation and/or activating NK (natural killer) cell proliferation of
XX CC and/or inducing NK toxicity in a mammal which involves administration of
XX CC K12 protein. It is also used for treating HIV-1 infection, cancer (T cell
XX CC leukaemia, acute lymphocytic leukaemia, cutaneous T cell lymphoma),
XX CC bacterial and viral infections, mediated by CD7. In the case of treating
XX CC T cell leukaemia the soluble K12 protein is covalently attached to a
XX CC toxin. A disease mediated by CD7 such as sepsis, graft versus host
XX CC disease due to transplantation, autoimmune diseases, multiple sclerosis,
XX CC arthritis, rheumatoid arthritis, psoriatic arthritis, scleroderma, lupus,
XX CC psoriasis, atopic dermatitis, type I diabetes mellitus, Hashimoto's
XX CC thyroiditis, pernicious anaemia, Addison's disease, myasthenia gravis,
XX CC lupus erythematosus and dermatomyositis, asthma, eczema, atopic
XX CC dermatitis, contact dermatitis, other eczematous dermatitis, seborrheic
XX CC dermatitis, and rhinitis is also treated by administering a K12
XX CC antagonist (neutralising antibody). The present sequence represents the
XX CC mouse K12 protein, which is given in the exemplification of the present
XX CC invention
XX SQ Sequence 212 AA;
    Query Match 24.9%; Score 326; DB 4; Length 212;
    Best Local Similarity 37.8%; Pred. No. 1.2e-26;
    Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
QY 10 GHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSWGENTVMSCNISNAFSHVNILKLR 69
Db 10 GLFPRMLWALLLLAASLSAQNHNVDPCCTEHEVSVMRSGSVVMACNISNLRDVTIELV 69
QY 70 AHQGESAIENEVAPGYFSRDGMQLOVGGVQAOLVTKGARDSHAGLYMWHLYHQHNNRQV 129
Db 70 TSEKTSIIFNHTPPGNSKDSWQLHIQGVQAOLVITDAQDKHSGNSWKLHGFQAEFKNF 129
QY 130 TLEVSAGAEPSAPDTGFWPVP-----AVTAVFILLVALVMFAWYRCRCSSQ 176
Db 130 NLTVNAADROKTEDLPTVKPDKPTAVRTEVIIIAITIIITIGVFWY----- 182
QY 177 RREKKFFLEPQMKVA 192
Db 183 ----KQPPVAPQIQMS 194
RESULT 11
ADJ76167
ID ADJ76167 standard; protein; 212 AA.
XX AC ADJ76167;
XX DT 20-MAY-2004 (first entry)
XX DE Marker gene related amino acid sequence SEQ ID NO:1419.
XX KW bronchial asthma; chronic obstructive pulmonary disease;
XX KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
XX KW gene therapy; marker.

```

OS Mus musculus.
 PN EP1394274-A2.
 PD 03-MAR-2004.
 PP 04-AUG-2003; 2003EP-00254857.
 XX 06-AUG-2002; 2002JP-00229312.
 PR 20-MAR-2003; 2003JP-00077212.
 XX (GENO-) GENOX RES INC.
 XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuwara K;
 XX WPI; 2004-193155/19.
 XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
 PT comparing the expression level of a marker gene in a biological sample
 PT from a subject with the expression level of the gene in a sample from a
 PT healthy subject.
 XX Claim 16; SEQ ID NO 1419; 241pp; English.
 XX The present invention describes a method of testing for bronchial asthma
 CC or chronic obstructive pulmonary disease. The method comprises
 CC determining the expression level of a marker gene in a biological sample
 CC from a subject, comparing the expression level determined with the
 CC expression level of the marker gene in a biological sample from a healthy
 CC subject, and judging whether the subject has bronchial asthma or chronic
 CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
 CC genes (S1) whose expression levels increase when respiratory epithelial
 CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
 CC whose expression levels decrease when respiratory epithelial cells are
 CC stimulated with interleukin-13. Also described: (1) a reagent (1) for
 CC testing for bronchial asthma or chronic obstructive pulmonary disease;
 CC (2) a kit for screening for a candidate compound for a therapeutic agent
 CC to treat bronchial asthma or chronic obstructive pulmonary disease;
 CC an animal model for bronchial asthma or chronic obstructive pulmonary
 CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
 CC method for producing an animal model for bronchial asthma or chronic
 CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
 CC asthma or chronic obstructive pulmonary disease, comprising the compound,
 CC a marker gene or an antisense nucleic acid corresponding to a portion of
 CC the marker gene, a ribozyme, a polynucleotide that suppresses the
 CC expression of the gene through an RNAi effect or an antibody recognising
 CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
 CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
 CC probe has been immobilised to assay a marker gene. (1) has respiratory
 CC and antiasthmatic activities, and can be used in gene therapy. The method
 CC is useful for testing for or screening for a therapeutic agent for
 CC bronchial asthma or chronic obstructive pulmonary disease. The present
 CC sequence is used in the exemplification of the present invention.
 XX SQ Sequence 212 AA;
 Query Match 24.9%; Score 326; DB 8; Length 212;
 Best Local Similarity 37.8%; Pred. No. 1.2e-26;
 Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
 QY 10 GHVSQALGTLFLAASLSAQNEGSDSPICTEGVVSVSGNNTVMSCNMFASHVNIKL 69
 DB 10 GLFPRMLWALLLLAASLNHNDVDEPCCTEHEVSVMNRGRVVMACNINSLRDVTIELV 69
 QY 70 AHQGESALFNEVAPGYFRDQWOLQVGGVLAOLVTKGARDSHAGLYVMHVLVGHQRNRQV 129
 DB 70 TSEKTSIIFNTTPGNSKDSNQLHIQVQALVITDAQDGHSGNSYKWLHGFQAEKFN 129
 QY 130 TLEVGASBPQAPDGFVPV-----AVVTAVFILLVALVMPFAWYRCRCSQ 176
 DB 130 NUTVNAARDQKTEDLPVTKVDPKPTAVRTEVIIIIATITTIITGIGVFVWY----- 182
 QY 177 REKKFFLLEPQMKVA 192

DB 183 ----KQFPVAPQIQMS 194
 RESULT 12
 ADJ76168
 ID ADJ76168 standard; protein; 212 AA.
 XX AC ADJ76168;
 XX DT 20-MAY-2004 (first entry)
 XX DE Marker gene related amino acid sequence SEQ ID NO:1420.
 XX KW bronchial asthma; chronic obstructive pulmonary disease;
 XX KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
 XX KW gene therapy; marker.
 XX OS Mus musculus.
 XX PN EP1394274-A2.
 XX PD 03-MAR-2004.
 XX 04-AUG-2003; 2003EP-00254857.
 XX 06-AUG-2002; 2002JP-00229312.
 XX 20-MAR-2003; 2003JP-00077212.
 XX (GENO-) GENOX RES INC.
 XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuwara K;
 XX WPI; 2004-193155/19.
 XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
 PT comparing the expression level of a marker gene in a biological sample
 PT from a subject with the expression level of the gene in a sample from a
 PT healthy subject.
 XX Claim 16; SEQ ID NO 1420; 241pp; English.
 XX The present invention describes a method of testing for bronchial asthma
 CC or chronic obstructive pulmonary disease. The method comprises
 CC determining the expression level of a marker gene in a biological sample
 CC from a subject, comparing the expression level determined with the
 CC expression level of the marker gene in a biological sample from a healthy
 CC subject, and judging whether the subject has bronchial asthma or chronic
 CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
 CC genes (S1) whose expression levels increase when respiratory epithelial
 CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
 CC whose expression levels decrease when respiratory epithelial cells are
 CC stimulated with interleukin-13. Also described: (1) a reagent (1) for
 CC testing for bronchial asthma or chronic obstructive pulmonary disease;
 CC (2) a kit for screening for a candidate compound for a therapeutic agent
 CC to treat bronchial asthma or chronic obstructive pulmonary disease;
 CC an animal model for bronchial asthma or chronic obstructive pulmonary
 CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
 CC method for producing an animal model for bronchial asthma or chronic
 CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
 CC asthma or chronic obstructive pulmonary disease, comprising the compound,
 CC a marker gene or an antisense nucleic acid corresponding to a portion of
 CC the marker gene, a ribozyme, a polynucleotide that suppresses the
 CC expression of the gene through an RNAi effect or an antibody recognising
 CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
 CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
 CC probe has been immobilised to assay a marker gene. (1) has respiratory
 CC and antiasthmatic activities, and can be used in gene therapy. The method
 CC is useful for testing for or screening for a therapeutic agent for
 CC bronchial asthma or chronic obstructive pulmonary disease. The present
 CC sequence is used in the exemplification of the present invention.
 XX SQ Sequence 212 AA;

The invention relates to human organelle-associated proteins (ORGA) and nucleic acid molecules encoding such proteins. ORGA sequences are useful for diagnosing, treating and/or preventing cell proliferative disorders (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers (or hepatitis), reproductive disorders (e.g. infertility, endometriosis, or ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal disorders (e.g. dysphagia, peptide oesophagitis, gastritis, indigestion, anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's disease, Pick's disease, Huntington's disease or dementia), urologic disorders (e.g. infection, urethritis, cystitis, prostatitis or upper tract infections) and renal disorders (e.g. glomerulonephritis, renal amyloidosis, renal failure, Addison's disease or hypertension). The invention is also useful in gene therapy. The present sequence is human ORGA protein

Query Match 22.5%; Score 294; DB 6; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.4e-23;
Matches 55; Conservative 0; Mismatches 0; Indels

Qy	1	MQTCLAPFGHVSAQLGTLILFLAASLSAQNEGDSPICTEGVSVSWGENTVMS	55
Dβ	1	MOTCPLAPFGHVSAQLGTLILFLAASLSAQNEGDSPICTEGVSVSWGENTVMS	55

AAB34671
ID AAB34671 standard; protein: 55 AA.

XX	
AC	AAB34671;
XX	
DT	26-JAN-2001 (first entry)

XX
DE
XX

Gene 35 human secreted protein homologous amino acid sequence #155.

Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
antitumatic; antiproliferative; cytostatic; cardiant; vasotropic;
cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm
hyperproliferative disorder; cancer; cardiovascular disorder; infection
cerebrovascular disorder; angiogenesis; nervous system disorder;
ocular disorder; wound healing; skin aging; food additive preservative

OS Homo sapiens.
XX
PN WO200056751-A

XX
PD 28-SEP-2000.

XX
PF
09-MAR-2000; 2000WO-US006013.

XX
PR 19-MAR-1999; 99US-0125360P.

PR 11-JUN-1999; 99US-0138626P.
PR 03-DEC-1999; 99US-0168662P.

XX PA (HUMA-) HUMAN GENOME SCI INC. 24

XX
PI
yy
Rosen CA, Ruben SM, Komatsoulis G;

WPI; 2000-579482/54.

Isolated nucleic acid molecule encoding a human secreted protein is used in preventing, treating or ameliorating a medical condition.

PS Disclosure: Page 412: 419pp: English: XX

xx The polynucleotide sequences given in AAC59738 to AAC59787 encode the
cc human secreted proteins given in AAB34577 to AAB34626. AAB34527 to
cc AAB34686 represent human secreted polypeptide sequences and proteins
cc homologous to them, which are given in the exemplification of the pre-
cc invention. Human secreted proteins have activities based on the tissue
cc and cells the genes are expressed in. Example of activities include:
cc

homologous to them, which are given in the exemplification of the present invention. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Example of activities include:

CC and cells the genes are expressed in. Example of activities include:

antiarthritic; immunosuppressive; antirheumatic; antiproliferative; cytoskeletal; cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide; fungicide; and ophthalmological. The polynucleotides and proteins can be used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases, hyperproliferative disorders e.g. neoplasms and cancers of the breast or liver, cardiovascular disorders, infections caused by disorders, angiogenesis, nervous system disorders. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The proteins can also be used as food additive or preservative to increase or decrease storage capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used in the exemplification of the present invention

XX Sequence 55 AA;

Query Match 21.9%; Score 286; DB 3; Length 55;
Best Local Similarity 100.0%; Pred. No. 3.9e-23;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 LRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQLFPYAADP 248
Db 1 LRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQLFPYAADP 55

RESULT 15
AAB34672
ID AAB34672 standard; protein; 55 AA.
XX AAB34672;
AC AAB34672;
XX 26-JAN-2001 (first entry)
XX Human secreted protein sequence encoded by gene 35 SEQ ID NO:156.
XX Human; secreted protein; diagnosis; immunosuppressive; antirheumatic; antiproliferative; cytoskeletal; cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm; hyperproliferative disorder; cancer; cardiovascular disorder; infection; cerebrovascular disorder; angiogenesis; nervous system disorder; ocular disorder; wound healing; skin aging; food additive; preservative.
XX Homo sapiens.
XX WO200056751-A1.
XX 28-SEP-2000.
XX 09-MAR-2000; 2000WO-US006013.
XX 19-MAR-1999; 99US-0125360P.
XX 11-JUN-1999; 99US-0138626P.
XX 03-DEC-1999; 99US-0168662P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM, Komatsoulis G;
XX WPI; 2000-579482/54.
XX Isolated nucleic acid molecule encoding a human secreted protein is used in preventing, treating or ameliorating a medical condition.
XX Disclosure; Page 412-413; 419pp; English.
XX The polynucleotide sequences given in AAC59738 to AAC59787 encode the

CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to AAB34686 represent human secreted polypeptide sequences and proteins homologous to them, which are given in the exemplification of the present invention. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Example of activities include: antiarthritic; immunosuppressive; antirheumatic; antiproliferative; cytoskeletal; cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide; fungicide; and ophthalmological. The polynucleotides and proteins can be used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases, hyperproliferative disorders e.g. neoplasms and cancers of the breast or liver, cardiovascular disorders, infections caused by disorders, angiogenesis, nervous system disorders. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The proteins can also be used as food additive or preservative to increase or decrease storage capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used in the exemplification of the present invention

XX Sequence 55 AA;

Query Match 21.3%; Score 279; DB 3; Length 55;
Best Local Similarity 98.2%; Pred. No. 2.3e-22;
Matches 54; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 194 LRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQLFPYAADP 248
Db 1 LRAGAQQGLSRASAEIWTDPSEPTPRPLALVFKPSPLGALLELLSPQLFPYAADP 55

Search completed: March 20, 2005, 09:47:33
Job time : 91 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 09:25:08 ; Search time 27 seconds
(without alignments)
685.665 Million cell updates/sec

Title: US-10-080-522-1

Perfect score: 1308

Sequence: 1 MOTCLAPPGHVSQALGTL.....PLGALELLSPQLFPFYAADP 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCITUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1308	100.0	248	US-08-755-559-1	Sequence 1, Appli
2	1308	100.0	248	US-09-210-474-1	Sequence 1, Appli
3	1308	100.0	248	US-09-539-774-1	Sequence 1, Appli
4	1308	100.0	248	US-09-997-165-4	Sequence 4, Appli
5	1308	100.0	248	US-09-949-016-6297	Sequence 6297, Ap
6	1308	100.0	258	US-09-949-016-8807	Sequence 8807, Ap
7	326	24.9	212	US-09-997-165-8	Sequence 8, Appli
8	90.5	6.9	1171	US-09-489-039A-10298	Sequence 10298, A
9	84.5	6.5	381	US-09-252-991A-27771	Sequence 27771, A
10	84.5	6.5	455	US-09-949-016-6949	Sequence 6949, Ap
11	84.5	6.5	455	US-09-949-016-11026	Sequence 11026, A
12	83.5	6.4	199	US-09-430-503-46	Sequence 46, Appl
13	83.5	6.4	199	US-09-430-503-48	Sequence 48, Appl
14	83.5	6.4	1912	US-09-949-016-10490	Sequence 10490, A
15	83	6.3	152	US-09-270-767-43788	Sequence 43788, A
16	82.5	6.3	288	US-08-396-650-1	Sequence 1, Appli
17	82.5	6.3	288	US-08-768-626-1	Sequence 1, Appli
18	82.5	6.3	288	US-09-645-069-12	Sequence 12, Appl
19	81.5	6.2	199	US-09-430-503-42	Sequence 42, Appl
20	81.5	6.2	199	US-09-430-503-44	Sequence 44, Appl
21	81.5	6.2	209	US-09-430-503-18	Sequence 18, Appl
22	81.5	6.2	209	US-09-430-503-20	Sequence 20, Appl
23	81.5	6.2	269	US-09-430-503-2	Sequence 2, Appli
24	81.5	6.2	269	US-09-430-503-4	Sequence 4, Appli
25	81.5	6.2	269	US-09-430-503-6	Sequence 6, Appli
26	81.5	6.2	269	US-09-430-503-8	Sequence 8, Appli
27	81.5	6.2	310	US-09-328-352-5485	Sequence 5485, Ap

28	81.5	6.2	313	4	US-09-949-016-10974	Sequence 10974, A
29	81.5	6.2	370	4	US-09-252-991A-27810	Sequence 27810, A
30	81	6.2	583	4	US-09-328-352-6422	Sequence 6422, Ap
31	80.5	6.2	209	3	US-09-430-503-22	Sequence 22, Appl
32	80.5	6.2	209	3	US-09-430-503-24	Sequence 24, Appl
33	80.5	6.2	310	4	US-09-949-016-8501	Sequence 8501, Ap
34	80.5	6.2	310	4	US-09-949-016-8896	Sequence 8896, Ap
35	79.5	6.1	343	4	US-09-252-991A-22974	Sequence 22974, A
36	79	6.0	205	4	US-09-949-016-9980	Sequence 9980, Ap
37	79	6.0	205	4	US-09-949-016-9981	Sequence 9981, Ap
38	79	6.0	205	4	US-09-949-016-9982	Sequence 9982, Ap
39	79	6.0	485	4	US-09-902-540-11019	Sequence 11019, A
40	79	6.0	545	4	US-10-237-551-121	Sequence 121, App
41	79	6.0	545	4	US-10-237-551-157	Sequence 157, App
42	79	6.0	547	4	US-10-237-551-216	Sequence 216, App
43	79	6.0	618	4	US-09-970-516-4	Sequence 4, Appli
44	79	6.0	618	4	US-09-817-676A-14	Sequence 14, Appl
45	78.5	6.0	225	4	US-09-530-139-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-08-755-559-1
; Sequence 1, Application US/08755559
; Patent No. 5912142
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTNY-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,559
; FILING DATE: 22-NOV-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-755-559-1

Query Match 100.0%; Score 1308; DB 2; Length 248;
Best Local Similarity 100.0%; Pred.No.2.6e-139;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MOTCLAPPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60
Db 1 MOTCLAPPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSCNISNA 60

us-10-080-522-1-rai

Sun Mar 20 14:17:42 2005

QY 61 FSHVNIKLRAHQESAIFNEVAPGYFSDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAIFNEVAPGYFSDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
 QY 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180
 DB 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180
 QY 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPGLGALELLSQP 240
 DB 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPGLGALELLSQP 240
 QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 2
 US-09-210-474-1
 ; Sequence 1, Application US/09210474
 ; Patent No. 6072034
 ; GENERAL INFORMATION:
 ; APPLICANT: KAUFMAN, RUSSEL E.
 ; APPLICANT: SLENTZ-KESLER, KIMBERLY
 ; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
 ; TITLE OF INVENTION: CELLS
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHVE P.C.
 ; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: U.S.A.
 ; ZIP: 22201-4714
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/210,474
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/755,559
 ; FILING DATE: 22-NOV-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WILSON, MARY J.
 ; REGISTRATION NUMBER: 32,955
 ; REFERENCE/DOCKET NUMBER: 1579-116
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 816-4000
 ; TELEFAX: (703) 816-4100
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 248 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-210-474-1

Query Match 100.0%; Score 1308; DB 3; Length 248;
 Best Local Similarity 100.0%; Pred. No. 2.6e-139;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
 DB 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
 QY 61 FSHVNIKLRAHQESAIFNEVAPGYFSDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAIFNEVAPGYFSDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120

QY 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180
 DB 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180
 QY 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPGLGALELLSQP 240
 DB 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPGLGALELLSQP 240
 QY 241 LFPYAADP 248
 DB 241 LFPYAADP 248

RESULT 3
 US-09-539-774-1
 ; Sequence 1, Application US/09539774
 ; Patent No. 6350615
 ; GENERAL INFORMATION:
 ; APPLICANT: KAUFMAN, RUSSEL E.
 ; APPLICANT: SLENTZ-KESLER, KIMBERLY
 ; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
 ; TITLE OF INVENTION: CELLS
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHVE P.C.
 ; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: U.S.A.
 ; ZIP: 22201-4714
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/539,774
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/210,474
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WILSON, MARY J.
 ; REGISTRATION NUMBER: 32,955
 ; REFERENCE/DOCKET NUMBER: 1579-116
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 816-4000
 ; TELEFAX: (703) 816-4100
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 248 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-539-774-1

Query Match 100.0%; Score 1308; DB 3; Length 248;
 Best Local Similarity 100.0%; Pred. No. 2.6e-139;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
 DB 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
 QY 61 FSHVNIKLRAHQESAIFNEVAPGYFSDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
 DB 61 FSHVNIKLRAHQESAIFNEVAPGYFSDGWLQVQGGVAQLVIKIGARDSHAGLYMHLV 120
 QY 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180
 DB 121 GHORNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVNFAYRCRCQORREK 180

Db 121 GHORNRQVTLVSGAEPOSADPTGFWPVPVAVVTAFFILLVALVMFAWYRCRCQQORREK 180
Qy 181 KFFLLEPOMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPGLGALLELLSQP 240
Db 181 KFFLLEPOMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPGLGALLELLSQP 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 4
US-09-997-165-4
; Sequence 4, Application US/09997165
; Patent No. 6762030
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fanslow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-997-165-4

Query Match 100.0%; Score 1308; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 2.6e-139;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MOTCLAPFGHVSQLGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Db 1 MOTCLAPFGHVSQLGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Qy 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGWLQVQGGVAQLVIKARDSHAGLYMHLV 120
Db 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGWLQVQGGVAQLVIKARDSHAGLYMHLV 120
Qy 121 GHORNRQVTLVSGAEPOSADPTGFWPVPVAVVTAFFILLVALVMFAWYRCRCQQORREK 180
Db 121 GHORNRQVTLVSGAEPOSADPTGFWPVPVAVVTAFFILLVALVMFAWYRCRCQQORREK 180
Qy 181 KFFLLEPOMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPGLGALLELLSQP 240
Db 181 KFFLLEPOMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPGLGALLELLSQP 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 5
US-09-949-016-6297
; Sequence 6297, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6297
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6297

Query Match 100.0%; Score 1308; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 2.6e-139;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MOTCLAPFGHVSQLGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Db 1 MOTCLAPFGHVSQLGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Qy 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGWLQVQGGVAQLVIKARDSHAGLYMHLV 120
Db 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGWLQVQGGVAQLVIKARDSHAGLYMHLV 120
Qy 121 GHORNRQVTLVSGAEPOSADPTGFWPVPVAVVTAFFILLVALVMFAWYRCRCQQORREK 180
Db 121 GHORNRQVTLVSGAEPOSADPTGFWPVPVAVVTAFFILLVALVMFAWYRCRCQQORREK 180
Qy 181 KFFLLEPOMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPGLGALLELLSQP 240
Db 181 KFFLLEPOMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPGLGALLELLSQP 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 6
US-09-949-016-8807
; Sequence 8807, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8807
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8807

Query Match 100.0%; Score 1308; DB 4; Length 258;
Best Local Similarity 100.0%; Pred. No. 2.7e-139;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MOTCLAPFGHVSQLGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Db 11 MOTCLAPFGHVSQLGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 70
Qy 61 FSHVNIKLRAHQESAIENEVAPGYFSRDGWLQVQGGVAQLVIKARDSHAGLYMHLV 120
Db 71 FSHVNIKLRAHQESAIENEVAPGYFSRDGWLQVQGGVAQLVIKARDSHAGLYMHLV 130

QY 121 GHORNRQVTLVSGABPOSAPDTGFWPVPVAVTAVFILLVALVWFAYRCRCQQORREK 180
DB 131 GHORNRQVTLVSGABPOSAPDTGFWPVPVAVTAVFILLVALVWFAYRCRCQQORREK 190
QY 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWPDPSEPTPRPLALVFKPSPIGALELLSPQP 240
DB 191 KFFLEPQMKVAALRAGAQQGLSRASAEIWPDPSEPTPRPLALVFKPSPIGALELLSPQP 250
QY 241 LFPYAADP 248
DB 251 LFPYAADP 258

RESULT 7
US-09-997-165-8
; Sequence 8, Application US/09997165
; Patent No. 6762030
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fonglow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-997-165-8

Query Match 24.9%; Score 326; DB 4; Length 212;
Best Local Similarity 37.8%; Pred. No. 1.9e-28;
Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
QY 10 GHVSQALGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSGNISNAPSHVNIKUR 69
DB 10 GLFPRMLALLLSALNAHNDVDEPCCTEHEVSVNRGVRVVMACNISNLRDVTIELV 69
QY 70 AHGQESAIFNEVAPGFSRDGWLQVGGVAQLVKGARDSHAGLYMHLVGHQRNRQV 129
DB 70 TSEKTSIIFNHTPPGNYSKDSWLHIOGVQVQLVITDAQDKHSGNYSWKLHGFQAEKFN 129
QY 130 TLEVSGAEPOSAPDTGFWPVPV-----AVVTAVFILLVALVWFAYRCRCQQO 176
DB 130 NLTVNAADRKTELDLPVKVDPKPTAVRTEVIIIAIAITIIITGIGVFWY----- 182
QY 177 RREKKFLLPEQMKVA 192
DB 183 ----KQPPVAPQIQMS 194

RESULT 8
US-09-489-039A-10298
; Sequence 10298, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10298
; LENGTH: 1171

; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10298

Query Match 6.9%; Score 90.5; DB 4; Length 1171;
Best Local Similarity 22.7%; Pred. No. 1.1;
Matches 57; Conservative 37; Mismatches 96; Indels 61; Gaps 13;
QY 19 LLFLAASLSAQNEGWDSPICTEGVSV-----VSWGENTV-----MSCNISNAPSHV 64
DB 582 LAFYAASL-ASHPQWRLLPV-DDGLVSVQVRLIRQLGQBNSESTLYQKLAQVANQYADM 639
QY 65 NI-KLRAHQESAIIF--NEVAPGFSRDGWLQVGGVAQLVKGARDSHAGLYMHLV 121
DB 640 RIADMTADTDASRLFTSTDEVFQWETQAWAQVPAIEKVVAE--RRDE---MDWVLS 694
QY 122 HORNROVT-----LEVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVWFAYRCRCQQO 177
DB 695 TKQTAAQSTSPALRAERLAEYFADFSGAW-----LDFLSLRWQRAATLSA 742
QY 178 REKKFLLPEQMK--VAALRAGAQQO-----LSRASAEIWPDPSEP----- 216
DB 743 IDQLTLMADVRSPLVALMNTLSVQGRGTGTGCEAIADSLVKSARQLFNRDNPVIDQRSG 802
QY 217 TPRPLALVFKP 227
DB 803 ARGPLDATFGP 813

RESULT 9
US-09-252-991A-27771
; Sequence 27771, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27771
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27771

Query Match 6.5%; Score 84.5; DB 4; Length 381;
Best Local Similarity 26.2%; Pred. No. 0.91;
Matches 60; Conservative 27; Mismatches 91; Indels 51; Gaps 12;
QY 1 MCTCLAPFGHVQSQAIGTLLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSGNISNA 60
DB 192 MVTAPLVMPVEITGLSLLLFVA---MAQLGWQP---ERGIVTI-WIAHT-----SFC 238
QY 61 FSHVNIKLRAHQESAIIFNEVAPGFSRDGWLQVGGVAQLVKGARDSHAGLYM 117
DB 239 SSVVAVVVSARLRELDLSIEEAAMDGLGAPKPKVFLITIPMIAPSLAAGWMSFA-LSLD 297
QY 118 HLHGQRNRQVTLVSG-----AEQSPADTGFVPVAVTAVFILLVAL-VMFAY 169
DB 298 DLV-----LASFVSGPGSTTLPEVFSVRLGVKPEINAVASLLISVSLFTFFAWY 349
QY 170 RCRCSQORREKKFFLLEPQMKVAALRAGAQQGLSRASAEIWPDPSEP 218
DB 350 FTRQAEER-----RRAIQAMEETATD-WQKGS-PTP 380

RESULT 10

Search completed: March 20, 2005, 09:50:12
Job time : 29 secs

Query Match	6.4%;	Score 83.5;	DB 4;	Length 1912;
Best Local Similarity	20.7%;	Pred. No. 14;		
Matches	62;	Conservative	31;	Mismatches 90;
				Indels 117; Gaps 15;
Qy	5	PLA	PPG	-HVSQALGTLFLAASLSAQNEG-----DSPICTEGVVSVW-- 47
Db	592	PL	PPG	TDHITVPL-ALMFEDVTVAATNFSGYDCSAVOALEAAAPCRACVGIWRCHWCP 650
Qy	48	-----	GENTVM	SCNITSNARSHVNIKLRAGHQSIAINEVAPGVFSRDGWOLQ 94
Db	651	QSSHC	VYGEHC	PEGERTIYSA-----QEVDDIQVGFAGCPQVGLAGP-HLVPVGWESH 703
Qy	95	VOG	GAQ	LVIKGARDSHAGLYMHVLGHQRNNRQVTLVSGAEPQSPADTGF-- 146
Db	704	LALVR	NRL--QHFR	GLPASFCWLELPCELRLPATLE-----ETAGDSGLIHCOAHQF 755
Qy	147	-----	WPYPAVVT	-----AVFTLLVALVMFMYRCK-----CSQORREKFF 182
Db	756	YPS	MSQ	RELFPVIYVYTOEAQRDLNTHALYVL-----YDCAMGHPDCSH-- 800
Qy	183	FLLE	FQMK	VAALRAGAQQGLSRASAEMLTWPSEPTPR--PLALVFKPSPLGAELELSPQ 240
Db	801	-----	COAANR	SLGCLWCADQCAPCYGPLIC-----PPGAVELLCFAP 838

RESULT 15
US-09-270-767-43788
; Sequence 43788, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 09:49:19 ; Search time 74 Seconds
(without alignments)
1107.760 Million cell updates/sec

Title: US-10-080-522-1
Perfect score: 1308
Sequence: 1 MQTCPLAPGHVSQALGTL.....PLGALLELLSQPLFPFYAADP 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1401741 seqs, 330541175 residues

Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1308	100.0	248	9 US-09-799-777-32	Sequence 32, Appl
2	1308	100.0	248	9 US-09-997-165-4	Sequence 4, Appl
3	1308	100.0	248	14 US-10-080-522-1	Sequence 1, Appl
4	326	24.9	212	9 US-09-997-165-8	Sequence 8, Appl
5	98	7.5	969	15 US-10-052-648A-35	Sequence 35, Appl
6	96	7.3	969	15 US-10-052-648A-34	Sequence 34, Appl
7	95.5	7.3	506	15 US-10-072-012-446	Sequence 446, App
8	91.5	7.0	597	15 US-10-425-114-62807	Sequence 62807, A
9	91	7.0	6620	15 US-10-080-334-290	Sequence 290, App
10	91	7.0	6620	15 US-10-408-765A-2291	Sequence 2291, Ap
11	91	7.0	7968	13 US-10-077-130-5	Sequence 5, Appl
12	89	6.8	435	10 US-09-759-130B-275	Sequence 275, App
13	89	6.8	435	14 US-10-189-123-5	Sequence 5, Appl

14	89	6.8	435	14	US-10-188-495-5	Sequence 5, Appl
15	89	6.8	435	16	US-10-741-790-275	Sequence 275, App
16	89	6.8	455	10	US-09-759-130B-273	Sequence 273, App
17	89	6.8	455	14	US-10-189-123-3	Sequence 3, Appl
18	89	6.8	455	14	US-10-188-495-3	Sequence 3, Appl
19	89	6.8	455	16	US-10-741-790-273	Sequence 273, App
20	84.5	6.5	271	14	US-10-156-761-14023	Sequence 14023, A
21	84.5	6.5	289	15	US-10-282-122A-43635	Sequence 43635, A
22	84.5	6.5	455	16	US-10-322-696-24	Sequence 24, Appl
23	83.5	6.4	199	14	US-10-095-131A-46	Sequence 46, Appl
24	83.5	6.4	199	14	US-10-095-131A-48	Sequence 48, Appl
25	83.5	6.4	366	14	US-10-091-438-153	Sequence 153, App
26	83.5	6.4	368	9	US-09-764-853-678	Sequence 678, App
27	83.5	6.4	468	16	US-10-626-832-19	Sequence 19, Appl
28	83.5	6.4	699	15	US-10-369-493-13729	Sequence 13729, A
29	82.5	6.3	286	16	US-10-367-094-107	Sequence 107, App
30	82.5	6.3	288	13	US-10-068-215-12	Sequence 12, Appl
31	82.5	6.3	288	14	US-10-115-615-12	Sequence 12, Appl
32	82.5	6.3	510	15	US-10-190-115-48	Sequence 48, Appl
33	82.5	6.3	510	15	US-10-190-115-135	Sequence 135, App
34	82.5	6.3	510	15	US-10-369-072-48	Sequence 48, Appl
35	82.5	6.3	510	15	US-10-042-865-59	Sequence 59, Appl
36	82.5	6.3	510	15	US-10-042-865-60	Sequence 60, Appl
37	82.5	6.3	785	9	US-09-989-920-218	Sequence 218, App
38	82.5	6.3	785	11	US-09-989-890-244	Sequence 244, App
39	82.5	6.3	805	15	US-10-042-865-4	Sequence 4, Appl
40	82.5	6.3	833	14	US-10-149-819-4	Sequence 4, Appl
41	82.5	6.3	833	15	US-10-190-115-14	Sequence 14, Appl
42	82.5	6.3	833	15	US-10-190-115-16	Sequence 16, Appl
43	82.5	6.3	833	15	US-10-190-115-18	Sequence 18, Appl
44	82.5	6.3	833	15	US-10-369-072-14	Sequence 14, Appl
45	82.5	6.3	833	15	US-10-369-072-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-799-777-32
; Sequence 32, Application US/09799777
; Patent No. US20020091244A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Hillman, Jennifer L.
; Corley, Neil C.
; Guegler, Karl J.
; Baugh, Mariah
; Sather, Susan
; Shah, Purvi
; TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS
; NUMBER OF SEQUENCES: 154
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/799,777
; FILING DATE: 06-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/002,485
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, LUCY J.
; REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0459 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 248 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 32 :
US-09-799-777-32

Query Match 100.0%; Score 1308; DB 9; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.4e-124;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Db 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGDARDSHAGLYMHLV 120
Db 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGDARDSHAGLYMHLV 120
QY 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVWFAYRCRCQORREK 180
Db 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVWFAYRCRCQORREK 180
QY 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPGLALELLSQP 240
Db 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPGLALELLSQP 240
QY 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 2
US-09-997-165-4
Sequence 4, Application US/09997165
Patent No. US20020141999A1
GENERAL INFORMATION:
APPLICANT: Lyman, Stewart D.
APPLICANT: Fanslow, William C.
TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
FILE REFERENCE: 2913-US
CURRENT APPLICATION NUMBER: US/09/997,165
CURRENT FILING DATE: 2001-11-27
PRIOR APPLICATION NUMBER: PCT/US00/14612
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/136,450
PRIOR FILING DATE: 1999-05-28
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 248
TYPE: PRT
ORGANISM: Homo sapiens
US-09-997-165-4

Query Match 100.0%; Score 1308; DB 9; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.4e-124;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Db 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGDARDSHAGLYMHLV 120
Db 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGDARDSHAGLYMHLV 120

QY 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVWFAYRCRCQORREK 180
Db 121 GHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILLVALVWFAYRCRCQORREK 180
QY 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPGLALELLSQP 240
Db 181 KFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTPRPLALVFKPSPGLALELLSQP 240
QY 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 3
US-10-080-522-1
Sequence 1, Application US/10080522
Publication No. US20030096326A1
GENERAL INFORMATION:
APPLICANT: KAUFMAN, RUSSEL E.
APPLICANT: SLENTZ-KESLER, KIMBERLY
TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER CELLS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/080,522
FILING DATE: 25-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/539,774
FILING DATE: 31-MAR-2000
APPLICATION NUMBER: US 09/210,474
FILING DATE: 14-DEC-1998
APPLICATION NUMBER: US 08/755,559
FILING DATE: 22-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-645
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 248 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-080-522-1

Query Match 100.0%; Score 1308; DB 14; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.4e-124;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
Db 1 MOTCPAPFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVSVSWGENTVMSNCISNA 60
QY 61 FSHVNIKLRAHQESAI FNEVAPGYFSRDGWLQVGGVAQLVIKGDARDSHAGLYMHLV 120

Db 61 FSHVNIKURAHGQSAIFNEVAPGYFSRSGWQLOVQGGVAQLVTKGARDSHAGLYMMHLV 120
Qy 121 QHQRNRQVTLVSGAEPQSPADTGFVFPVAVTAVFILLVALVWFAYRCRCQQORREK 180
Db 121 QHQRNRQVTLVSGAEPQSPADTGFVFPVAVTAVFILLVALVWFAYRCRCQQORREK 180
Qy 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTPTDSEPTPRPLALVFKSPIGALBELLSPQ 240
Db 181 KFFLEPQMKVAALRAGAQQGLSRASAEIWTPTDSEPTPRPLALVFKSPIGALBELLSPQ 240
Qy 241 LFPYAADP 248
Db 241 LFPYAADP 248

RESULT 4
US-09-997-165-8
; Sequence 8, Application US/09997165
; Patent No. US20020141999A1
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fanslow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Mus sp.
; ORGANISM: Mus sp.
US-09-997-165-8

Query Match 24.9%; Score 326; DB 9; Length 212;
Best Local Similarity 37.8%; Pred. No. 1.1e-24;
Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
Qy 10 GHVSQALGTLFLAASLAQNEGWDSPTCTEGVSVSWGENTVMSCNISNAFVSHVNIKLR 69
Db 10 GLFPRMLWALLLAASLAHNDVMDPEPCCTHEVSVNRGRVVMACNISNLRDVTIELV 69
Qy 70 AHGQESAIFNEVAPGYFSRSGWQLOVQGGVAQLVTKGARDSHAGLYMMHLVGHQRNRQV 129
Db 70 TSEKTSIIFNHTPPGNSKDSQNLHIQVQALVITDAQDKHSGNYSWKLHGFQAEFKNF 129
Qy 130 TLEVSGAEPQSPADTGFVFPV-----AVTAVFILLVALVWFAYRCRCQQ 176
Db 130 NLTVNAADROKTEDLPVTKVPDKPTAVRTEVIIIIATITIIITIGVFWY----- 182
Qy 177 RREKKFFLEPQMKVA 192
Db 183 ----KQFPVAPQIQMS 194

RESULT 5
US-10-052-648A-35
; Sequence 35, Application US/10052648A
; Publication No. US2004000558A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steven
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gunther, Erik
; APPLICANT: Kekuda, Ramesh

; APPLICANT: MacDougall, John R.
; APPLICANT: Mehraban, Fuad
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Shimkets, Richard
; APPLICANT: Smithson, Glennda
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-250 (CURA-550)
; CURRENT APPLICATION NUMBER: US/10/052,648A
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,920
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/284,549
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/303,229
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,605
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/269,098
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/264,159
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 969
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (848)..(889)
; OTHER INFORMATION: Where Xaa is any amino acid
US-10-052-648A-35
Query Match 7.5%; Score 98; DB 15; Length 969;
Best Local Similarity 24.9%; Pred. No. 1.4;
Matches 51; Conservative 24; Mismatches 86; Indels 44; Gaps 10;
Qy 2 QTCPLAFPG---HVSQALGTLFLAASLSAQN-----EGWDSPTCTEGVSVSWGENT 51
Db 620 QACPPGFWGPAFCFACSCHN-----GASCSAEDGACHCTPGWTGLFQRCPPAFAFGKDC 674
Qy 52 VMSCNISNAFS--HVNIKLRH-----QGESAIFNEVAPGYFSRSGWQ-----QV 95
Db 675 GRVCCQNGASCDHISGKCTCTGTGHCHE--QRCAPGTGCGCQQLCECMNNSTCDHV 732
Qy 96 QGG--VAQLVTKGARDSHAGLYMMHLVGHQRNRQVTLVSGAEPQSPADTGFVFPVAVT 154
Db 733 TGTCCYSPGFGKIRCDQAALMMEELNPNYTKISPAL-----GAERHSV-----GAVTGMIL 782
Qy 155 AVFILLVALVWFAYRCRCQQORRE 179
Db 783 LLFLIVVLLGLFAWHRRRRQKEKGD 807
RESULT 6
US-10-052-648A-34
; Sequence 34, Application US/10052648A
; Publication No. US2004000558A1

GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steven
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gunther, Erik
; APPLICANT: Kekuda, Ramesh
; APPLICANT: MacDougall, John R.
; APPLICANT: Mehraban, Fuad
; APPLICANT: Patturajan, Meera
; APPLICANT: Rottenberg, Mark
; APPLICANT: Shinkets, Richard
; APPLICANT: Smithson, Glenda
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-250 (CURA-550)
; CURRENT APPLICATION NUMBER: US/10/052,648A
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,920
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/284,549
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/303,229
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,605
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/269,098
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/264,159
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 969
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-052-648A-34

Query Match 7.3%; Score 96; DB 15; Length 969;
Best Local Similarity 24.9%; Pred. No. 2.2;
Matches 51; Conservative 23; Mismatches 87; Indels 44; Gaps 10;

QY 2 QTCPLAFPG----HVSQALGTLFLAASLSAQN-----EGWDSPICTEGVSVSGENT 51
DB 620 QACPPGFWGPACFHACSCN-----GASCSAEDGACHCTPGTGLFCTQRCPPAFAFGKDC 674

QY 52 VMSCNISNAFS--HVNKILRAH-----GQSAIFNEVAPGFYSRDQWL-----QV 95
DB 675 GRVCOQNGASCDHISGKCTKCTGTGTGQHC--QRCAPGTGYGQQLCECMNNSTCDHV 732

QY 96 QGG--VAQLVIKGARDSHAGLYNMHLVGHORNNRQVTEVSGAEFQSPADPTGFWVPVAVT 154
DB 733 TGTCTCSPGFGIRCDQQAALNMEELNPTYTKISPAL-----GAERHSV-----GAVTGIML 782

QY 155 AVFILLVALVMEAWYRCRCQORRE 179
DB 783 LLFFIVLLGLFAWHRRRQKRGDRD 807

RESULT 7
US-10-072-012-446
; Sequence 446, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 446
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-072-012-446

Query Match 7.3%; Score 95.5; DB 15; Length 506;
Best Local Similarity 26.6%; Pred. No. 1;
Matches 51; Conservative 25; Mismatches 79; Indels 37; Gaps 8;

QY 4 CPLAFPGHVSQALGTLFLAASLSAQN-----EGWDSPICTEGVSVSGENTVMSCNISNAF-S 62
DB 228 CEVA---HVTLQGGPPLRGNTANLS---ETIRVPPTLEITRSPSAGNQVNVTCQVNFYPR 281

QY 63 HVNKLRAHG-----QESAIFNEVAPGFYSRDQWLQVGGVAQLVIKGARDSHAGLYNM 117

Db 282 HLQLTWLENGNMRTEAASVLVENKQGTFTQNTSMLLVNSSAHREAVLTCQ----- 332
Qy 118 HLUGHQRN---NRQVTLVSGAEPOSADPTGFWPVP-----AVVTAVFILLVALVM 165
Db 333 ---VEHDROPVSKNHTLEVSA--PQXQDQGTGTCFNDNNWTSIFIVGVVVCALLVALLI 388
Qy 166 FAWYRCRCSQOR 177
Db 389 AALYLLRIRQNK 400

RESULT 8
US-10-425-114-62807
; Sequence 62807, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21153313B
; CURRENT APPLICATION NUMBER: US/10/425.114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62807
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3279-208-G7_FLI.pep
US-10-425-114-62807

Query Match 7.0%; Score 91.5; DB 15; Length 597;
Best Local Similarity 19.9%; Pred. No. 3.3;
Matches 62; Conservative 44; Mismatches 84; Indels 121; Gaps 16;

Qy 21 FLAASLSAQNGWSP-----ICTEGVSVSW-----GE 49
Db 299 FKAASAAAGHPDWELFDAGINDPTDGTFTTAERGTYLTEQGRFFLTWYSRKLIOHGD 358
Qy 50 NTVMSCSNISAFSHVNIKLRA-----HGQESALFNEVAPGYFS---RDGQWLQVQ 96
Db 359 RVLDEAN--KAPLGCKVKLAAKVSGIHWYRHPSHAA---ELTSGYNNLGGRDGY----- 408
Qy 97 GGVAQLVTKGARDSHAGLYMHLVGHQRNRRQVTLVSGAEPQAPDPTGFWPVPVAVTAV 156
Db 409 APIARMLAR-----HDGAVLNFTCAEMNSQ-----ABEALSAPEQ----- 445
Qy 157 FILLVALVYFAWYR-----CRCQORREKFF---LLEPQMKVAALBAGAQOGLS----- 203
Db 446 ---LVQVLSAGWREGVEVACENALSYDRRGYNQMLNARPNVGLSGAARRRVAATF 502
Qy 204 -RASAEEL-----WTPDSEPTPRPLALVFKPSPGLGALELL-----S 237
Db 503 LRLSDELLASNNFRIFRIFVRKMHADLDYCPDADRYGRPLKPLERSAPEMPERLLEATA 562
Qy 238 POPLFFPYAADP 248
Db 563 PAPAFPPF--DP 571

Query Match 7.0%; Score 91; DB 15; Length 6620;
Best Local Similarity 22.8%; Pred. No. 92;
Matches 58; Conservative 34; Mismatches 102; Indels 60; Gaps 10;

Qy 11 HVSQLGTLLFLAASLSAQNGWSPICTEGV-----VSVWCGENTVMSCNISNA 60
Db 1779 HGAQVLDLSAIY---SCRVGAEQDFPQVVEVAAKFCELLEPVGELGGTVTTLACELSPA 1835
Qy 61 FSHVNIKLRAHQBSAIFNEVAPGYFSRDGQWLQVQGVQVAFKGRDSDHAGLYMHLV 120

Query Match 7.0%; Score 91; DB 15; Length 6620;
Best Local Similarity 22.8%; Pred. No. 92;
Matches 58; Conservative 34; Mismatches 102; Indels 60; Gaps 10;

Qy 11 HVSQLGTLLFLAASLSAQNGWSPICTEGV-----VSVWCGENTVMSCNISNA 60
Db 1779 HGAQVLDLSAIY---SCRVGAEQDFPQVVEVAAKFCELLEPVGELGGTVTTLACELSPA 1835
Qy 61 FSHVNIKLRAHQBSAIFNEVAPGYFSRDGQWLQVQGVQVAFKGRDSDHAGLYMHLV 120

Sun Mar 20 14:17:43 2005

TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
TITLE OF INVENTION: Members and Uses Therefor
FILE REFERENCE: MPI2001-047P1RCF1(M)
CURRENT APPLICATION NUMBER: US/10/077,130
CURRENT FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: 60/269201
PRIOR FILING DATE: 2001-02-15
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 7968
TYPE: PRT
ORGANISM: Homo sapiens
US-10-077-130-5

Db 1836 CAEVVWRC-----GNTQPRVGR--FQVAGPVSRLTVGLRAEDAGEY---V 1879
Qy 121 GHORNNR---QVTLV-----SCAEPQAPDTGFWPVPVAVTAVFILLVALVMPAWYR 170
Db 1880 CESRDDHTSAQITVSVPRVWFMSGLSTVVAEGEGEATFCQVSPSDVAVV-----WFR 1933
Qy 171 CRCSQORREKFFLLEPQMKVAALRAGAQQGLSRASAEI-----WTPDSEPTPRPLALVF 225
Db 1934 DGA-----LLQPSKFAISQSGASHSLTISDLVLEDAGQITVEAGASSAALRV 1983
Qy 226 KPSPLGALELLSPQ 239
Db 1984 REAPVLFKKLEPQ 1997

Query Match 7.0%; Score 91; DB 16; Length 6620;
Best Local Similarity 22.8%; Pred. No. 92;
Matches 58; Conservative 34; Mismatches 102; Indels 60; Gaps 10;
RESULT 10
US-10-408-765A-2291
; Sequence 2291, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Wartock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2291
; LENGTH: 6620
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2291

Query Match 7.0%; Score 91; DB 16; Length 6620;
Best Local Similarity 22.8%; Pred. No. 92;
Matches 58; Conservative 34; Mismatches 102; Indels 60; Gaps 10;
RESULT 11
US-10-077-130-5
; Sequence 5, Application US/10077130
; Publication No. US20020168742A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
Qy 11 HVSQALGTLFLAASLSAQNEGWDSPTCTEGV-----VSVSGENTVMSCNISNA 60
Db 1779 HGAQVLDLSAIY---SCRVGAEGQDPFVQVEEVAAKFCRLLEPVSRLTVGLRAEDAGEY 1835
Qy 61 FSHVNIKLRAHQESAIFFNEVAPGYFSRDGWLQVQGGVAQLVIKGARDSHAGLYMHLV 120
Db 1836 CAEVVWRC-----GNTQPRVGR--FQVAGPVSRLTVGLRAEDAGEY---V 1879
Qy 121 GHORNNR---QVTLV-----SCAEPQAPDTGFWPVPVAVTAVFILLVALVMPAWYR 170
Db 1880 CESRDDHTSAQITVSVPRVWFMSGLSTVVAEGEGEATFCQVSPSDVAVV-----WFR 1933
Qy 171 CRCSQORREKFFLLEPQMKVAALRAGAQQGLSRASAEI-----WTPDSEPTPRPLALVF 225
Db 1934 DGA-----LLQPSKFAISQSGASHSLTISDLVLEDAGQITVEAGASSAALRV 1983
Qy 226 KPSPLGALELLSPQ 239
Db 1984 REAPVLFKKLEPQ 1997

Query Match 7.0%; Score 91; DB 13; Length 7968;
Best Local Similarity 22.8%; Pred. No. 1.2e+02;
Matches 58; Conservative 34; Mismatches 102; Indels 60; Gaps 10;
RESULT 12
US-09-759-130B-275
; Sequence 275, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: MPI00-5350NNIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14

Query Match	6.8%;	Score 89;	DB 14;	Length 435;
Best Local Similarity	31.0%;	Pred. No. 3.8;		
Matches	44;	Conservative 18;	Mismatches 62;	Indels 18; Gaps 9;

Qy	82	APGVFSRDGQHQVQGGVAQVLIKGRDSSAGLYMWHLVG---	HQRNNRQVTLVSVGAEP	138
Db	301	APG--SRDGSIAVLADG--SLAIGNVQEQHAGLVFVCLATPRLHNNQTHEYNVSVHPRP		356
Qy	139	Q-SAPDTGFVFPVPAVTVTAIVILLVYMFAYWYRCSCQQRREKFFLLEPQMKVAALRAG		197
Db	357	EPFAFNTGF--TTLLGCAVGLVLLVLYLFA--PPRCRC--RRACPPLPPLANTOPAP--RAE		410

```

1  GENERAL INFORMATION:
2  /
3  / APPLICANT: Millennium Pharmaceuticals, Inc.
4  /
5  / APPLICANT: McCarthy, Sean A
6  /
7  / APPLICANT: Frazer, Christopher C
8  /
9  / APPLICANT: Sharp, John D
10 /
11 / APPLICANT: Barnes, Thomas S
12 /
13 / APPLICANT: Kirt, Susan J
14 /
15 / APPLICANT: Mackay, Charles R
16 /
17 / APPLICANT: Myers, Paul S
18 /
19 / APPLICANT: Leiby, Kevin R
20 /
21 / APPLICANT: Wrighton, Nicolas
22 /
23 / APPLICANT: Goodearl, Andrew
24 /
25 / APPLICANT: Holzman, Douglas A
26 /
27 / TITLE OF INVENTION: NOVEL GENES ENCODING PR
28 /
29 / TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC
30 /
31 / TITLE OF INVENTION: USES.
32 /
33 / FILE REFERENCE: MP100-5350N1M
34 /
35 / CURRENT APPLICATION NUMBER: US/10/741,790
36 /
37 / CURRENT FILING DATE: 2003-12-19
38 /
39 / PRIOR APPLICATION NUMBER: US 09/479,249
40 /
41 / PRIOR FILING DATE: 2000-01-07

```

```

, PRIOR APPLICATION NUMBER: US 09/559,497
, PRIOR FILING DATE: 2000-04-27
, PRIOR APPLICATION NUMBER: US 09/578,063
, PRIOR FILING DATE: 2000-05-24
, PRIOR APPLICATION NUMBER: US 09/333,159
, PRIOR FILING DATE: 1999-06-14
, PRIOR APPLICATION NUMBER: US 09/596,194
, PRIOR FILING DATE: 2000-07-14
, PRIOR APPLICATION NUMBER: US 09/342,364
, PRIOR FILING DATE: 1999-06-29
, PRIOR APPLICATION NUMBER: US 09/608,452
, PRIOR FILING DATE: 2000-06-30
, PRIOR APPLICATION NUMBER: US 09/393,996
, PRIOR FILING DATE: 1999-09-10
, PRIOR APPLICATION NUMBER: US 09/602,871
, PRIOR FILING DATE: 2000-06-23
, PRIOR APPLICATION NUMBER: US 09/420,707
, PRIOR FILING DATE: 1999-10-19
, NUMBER OF SEQ ID NOS: 460
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 275
, LENGTH: 435
, TYPE: PRT
, ORGANISM: Homo sapiens
US-10-741-790-275

Query Match      6.8%; Score 89; DB 16; Length 435;
Best Local Similarity 31.0%; Pred. No. 3.8;
Matches 44; Conservative 18; Mismatches 62; Indels 18; Gaps 9;

QY      82  AFGYERDQWQLQVQGGAQVLTKGARDSDHAGLYMHLVG---HQRNNRQVTLVSGAEP 138
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      301  AFG--SRDGSIAVLADG--SLAIGNVOEQHAGLFVCLATGPRLHHNQTHEYNVSVHPPRP 356
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      139  Q-SAPDQFWPFAVTVAVFILLVALVFWAWYRCRQORREKKFLLIEPQMKVAALEAG 197
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      357  EEEAFNTGTF--TTLLGCAVGLVLLYLFA--PPCRCC--RRACPLPPLAPNTQAP--RAE 410
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      198  AQQGLSRASAEIWTDPDSEPTPR 219
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      411  PH-----KSVSLSTTPPDAPSQ 428
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: March 20, 2005, 10:15:10
Job time : 77 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using ew model

Run on: March 20, 2005, 09:17:28 ; Search time 25 Seconds
(without alignments)
954.470 Million cell updates/sec

Title: US-10-080-522-1
Perfect score: 1308
Sequence: 1 MOTCLAPFGHVSOALGTLL.....PLGALELLSQPLFFYAADP 248
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 9621673 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103.5	7.9	550	S64725	probable lipoprote
2	94.5	7.2	393	T07167	probable isocitrat
3	90.5	6.9	271	T36885	probable membrane
4	90	6.9	1088	H96747	unknown protein fl
5	89	6.8	798	A12053	competence protein
6	86.5	6.6	929	A44048	genome polypeptid
7	86	6.6	493	S51574	mocR protein - Rhi
8	85.5	6.5	420	A82699	conserved hypothet
9	85.5	6.5	468	B46114	glycoprotein gp13
10	85.5	6.5	2303	GNNTTP	genome polypeptid
11	85.5	6.5	2303	S13554	genome polypeptid
12	84.5	6.5	289	A83608	polyamine transpor
13	83.5	6.4	468	1 VGBEEH	glycoprotein gp13
14	83	6.3	550	1 VGBE18	glycoprotein E - h
15	83	6.3	436	E71086	hypothetical prote
16	82.5	6.3	233	T15620	hypothetical prote
17	82.5	6.3	288	A55737	PD-1 protein - hum
18	82.5	6.3	636	F72736	hypothetical prote
19	82	6.3	398	T46475	hypothetical prote
20	81	6.2	450	1 FOLJFP	gag polypeptid -
21	81	6.2	450	S23819	gag protein - feli
22	80.5	6.2	312	C87562	conserved hypothet
23	80.5	6.2	541	T48811	hypothetical prote
24	80.5	6.2	1240	T03097	CDO protein - huma
25	80.5	6.2	2301	1 GNNVTN	genome polypeptid
26	80.5	6.2	2303	1 GNNVTM	genome polypeptid
27	80	6.1	290	T09260	aquaporin-like tra
28	79	6.0	407	B70962	hypothetical prote
29	79	6.0	950	T51134	ionotropic glutama

30 78 6.0 204 2 T51012 hypothetical prote
31 77.5 5.9 390 1 Q0BE77 glycoprotein I pre
32 77.5 5.9 834 2 S66498 M-sema F protein p
33 77 5.9 99 2 C46518 Ig L1 chain V regi
34 77 5.9 124 2 S08640 hypothetical prote
35 77 5.9 287 2 S60455 transmembrane prot
36 77 5.9 289 2 S33617 trig-31 protein - g
37 77 5.9 303 2 A40807 membrane glycoprot
38 77 5.9 364 2 H75466 lipopolysaccharide
39 77 5.9 442 1 XUBRVS 3-phosphohikimate
40 77 5.9 910 2 D75524 alpha-dextran endo
41 77 5.9 3097 2 T00021 DN-cadherin - frui
42 77 5.9 26926 1 I38344 titin, cardiac mus
43 76.5 5.8 351 2 S39603 class I histocompa
44 76.5 5.8 666 2 H83943 transketolase tkt
45 76.5 5.8 835 2 S40140 ribonucleoside-dip

ALIGNMENTS

RESULT 1

S64725
probable lipoprotein uxpA precursor - Pseudomonas putida
C:Species: Pseudomonas putida
C>Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 09-Jul-2004
C:Accession: S64725; S47503
R:de Groot, A.; Krijger, J.J.; Filloux, A.; Tommassen, J.
Mol. Gen. Genet. 250, 491-504, 1996
A:Title: Characterization of type II protein secretion (xcp) genes in the plant growth-
A:Reference number: S64724; MUID:96186881; PMID:8602167
A:Accession: S64725
A:Molecule type: DNA
A:Residues: 1-550 <DEG>
A:Cross-references: UNIPROT:Q52289; EMBL:X81085; NID:G3293032; PIDN:CAA56977.1; PID:G53
C:Genetics:
A:Gene: uxpA
F:1-23/Domain: signal sequence #status predicted <SIG>
F:30-550/Product: probable lipoprotein uxpA #status predicted <MAT>

Query Match 7.9%; Score 103.5; DB 2; Length 550;
Best Local Similarity 23.2%; Pred. No. 0.17;
Matches 72; Conservative 42; Mismatches 118; Indels 79; Gaps 16;
QY 2 QTCPLAPFGHVSOA-LGTLLFLAA-----SLSAQN-EGWDSPTCTEGVVSVS 46
DB 90 QLAFLDAGRADEANLGGYAVLGALLEQLRGAGAGNSLTLENGQWNG-----SGLAYLT 145
QY 47 WQENTVMSGNISNAFSHVNIKLRA--HQGESAI FNEVAPGYFSR-----DGWQ 92
DB 146 QGESGVQSGQLLGSEARVSSDERVLWPQRSAAALYRQASATTILGAGLADQRQALGLEPLQ 205
QY 93 LOVQGVQVQLVTKG-----ARDSHAGLYMWH---LVGHQNNRVQVTLVSGAEPQAP-- 142
DB 206 LFERGG-ARIAVGVVTDPAQDQKASLKQWYQALLPVFOQARREADLVVALADVGTGPG 264
QY 143 -----DTGFVFPVAVVTAVFILLVALVMFAW-----YRCRC-----SQ 176
DB 265 WLAEPLPAIDLLLCARGQDLWTFPVQATQASGRVP-VLFACRGSGAFRLRCQVAGQW 323
QY 177 REEKXFF-----LLBFPQMKVAALRAGAQQGLSRASAEMLWTPDSEPTFR-PLALVFKPSPL 230
DB 324 QPEGFRFFPTFEQLTSLPAAQVRAGQALNQQRAGHAAML--DQPLARAPQALWRRDTRG 381
QY 231 GALELLSQPL 241
DB 382 GSWDRLLHQAL 392

RESULT 2

T07167
probable isocitrate dehydrogenase (NAD) (EC 1.1.1.41) chain 1 - tomato
C:Species: Lycopersicon esculentum (tomato)

Db 228 GVGKVTASPL-----PGPSGNVE 245

RESULT 4

H96747

unknown protein T10D10.14 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: H96747

R;Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; Chin, C.W.; Hughes, B.; Huizar, L.

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-393 <ANT>

A;Cross-references: UNIPROT:082004; EMBL:Y16126; PIDN:CAA76076.1

A;Experimental source: cultivar Early Mech; root

C;Superfamily: 3-isopropylmalate dehydrogenase

C;Keywords: oxidoreductase

Query Match 7.2%; Score 94.5; DB 2; Length 393;

Best Local Similarity 23.8%; Pred. No. 0.76; Indels 65; Gaps 9;

Matches 51; Conservative 30; Mismatches 68; Indels 65; Gaps 9;

QY 63 HVNKLRAHQESAINEVAPG---YFSRDGWLQVGGVAQLVIKAR--DSHAGL-Y 115

Db 94 HGMKEYASGDVDGIGN---PGRNKVFKEGKTPVGGVSSLVNQLRKELDLVSLVH 150

QY 116 MHWLVGHQRNRQVTL-----EVSAGAEQSPAP-----DT 144

Db 151 CFNLKGLPTRHENVDIVIRENTEGEYSGLHEHVPVGVESLKVMTKFCSERIAKYAFEY 210

QY 145 GFVPVAVTVAFV---ILLVALVFMFAYRCRCQQRRKKFFLLEPQMKVAALRAGAQQG 201

Db 211 AYLNRRKVVAVHAKANIMENLLMVFWSG-----RDRKQITLG----- 250

QY 202 LSRASAEIWPDSSETPRPLALVFKPSPLGAL 235

Db 251 -SSTMRELTTVACNLVQSLSLNMLMSPLISTEI 283

RESULT 3

T36885

probable membrane protein - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T36885

R;Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-271 <MUR>

A;Cross-references: UNIPROT:Q98218; EMBL:AL109848; PIDN:CAB52846.1; GSPDB:GN00070; SCORE

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCODDB:SC151.22C

C;Superfamily: hemolysin homolog yqxc

Query Match 6.9%; Score 90.5; DB 2; Length 271;

Best Local Similarity 25.1%; Pred. No. 1.1; Indels 71; Gaps 8;

Matches 51; Conservative 21; Mismatches 60; Indels 71; Gaps 8;

QY 83 PGYFSRDGWLQVGGVQAQ-----LVIKAR-----DSH 111

Db 63 PIVYVSGGHKL--AGALAAFPVHGLVVEGRRALDAGASTGGFTDVLRLAGAAHVAVVDVG 120

QY 112 AGLYMWHL-----VGHQRNRQVTLVSGAEFQSPADTGFVPVAVTVAVFILLVALV 164

Db 121 YGQLAWSLRQDRVTVKORTNRELTPEAIDG-----PVDLVGDLSTFPLALV 170

QY 165 MFAWYRCRCQQRRKKFFLLEPQMKV-----AALRAGAQQGLSRASAEIWT 211

Db 171 LPALVFC---TRPGADLVNMFQPEVGVKRLGSGGVVRSQAQLRAEAVRGVARKAWELGL 227

QY 212 PDSEPTPRPLALVFKPSPLGAL 234

Query Match 6.9%; Score 90; DB 2; Length 1088;

Best Local Similarity 21.2%; Pred. No. 6.3; Indels 66; Gaps 11;

Matches 55; Conservative 46; Mismatches 93; Indels 66; Gaps 11;

QY 15 ALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCTISNAPSHVNIKLRAHG-- 72

Db 789 ALGNQLNWAGR-GMGGTGISSSMSVPGIGN--QGNP-MNLNPASNLNAISQQLRSALT 844

QY 73 -QESAIFNEVAPGYFSRDGWLQVGGVAQLVIKARDSH---AGLYMWHLVGHQRNRQ 128

Db 845 PQNALFTQIRMGANRGVNGVAPQTGISG--VSGTRQMPHSSAGLSM----- 890

QY 129 VTLEVSGAEFQSPADTGFVPVAVTVAVFILLVALVFMFAYRCRCQQRRKKFFLLEPQ 188

Db 891 --LDQNRANLQRAAGMNGMGPFLMGMNLYM-----NQOQQOQLQQOPQ 935

QY 189 MKVAALRAGAQQGLSRASAEI-WTP-----DSEPTPRPLALVFKPSPL 230

Db 936 QQQLQHQQQLQPMSPSQQLAQSPQQOQLQHQHEPQQOQQOQATASPLQSVLSPPQV 995

QY 231 GA-----LELLSQPL 241

Db 996 GSPSAGITQQQLQQSSPQOM 1015

RESULT 5

AI2053

competence protein [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C;Accession: AI2053

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, DNA Res. 8, 2005-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AI2053

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-798 <KUR>

A;Cross-references: UNIPROT:Q9YV14; GB:BA000019; PIDN:BA073682.1; PID:g17131073; GSPDB

A;Experimental source: strain PCC 7120

C;Genetics:

Db 228 GVGKVTASPL-----PGPSGNVE 245

RESULT 4

H96747

unknown protein T10D10.14 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: H96747

R;Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; Chin, C.W.; Hughes, B.; Huizar, L.

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1088 <STO>

A;Cross-references: UNIPROT:Q9C9D7; GB:AE005173; MID:g6730762; PIDN:AAF27151.1; GSPDB:G

C;Genetics:

A;Gene: T10D10.14

A;Map position: 1

Query Match 6.9%; Score 90; DB 2; Length 1088;

Best Local Similarity 21.2%; Pred. No. 6.3; Indels 66; Gaps 11;

Matches 55; Conservative 46; Mismatches 93; Indels 66; Gaps 11;

QY 15 ALGTLFLAASLSAQNEGWDSPICTEGVSVSGENTVMSCTISNAPSHVNIKLRAHG-- 72

Db 789 ALGNQLNWAGR-GMGGTGISSSMSVPGIGN--QGNP-MNLNPASNLNAISQQLRSALT 844

QY 73 -QESAIFNEVAPGYFSRDGWLQVGGVAQLVIKARDSH---AGLYMWHLVGHQRNRQ 128

Db 845 PQNALFTQIRMGANRGVNGVAPQTGISG--VSGTRQMPHSSAGLSM----- 890

QY 129 VTLEVSGAEFQSPADTGFVPVAVTVAVFILLVALVFMFAYRCRCQQRRKKFFLLEPQ 188

Db 891 --LDQNRANLQRAAGMNGMGPFLMGMNLYM-----NQOQQOQLQQOPQ 935

QY 189 MKVAALRAGAQQGLSRASAEI-WTP-----DSEPTPRPLALVFKPSPL 230

Db 936 QQQLQHQQQLQPMSPSQQLAQSPQQOQLQHQHEPQQOQQOQATASPLQSVLSPPQV 995

QY 231 GA-----LELLSQPL 241

Db 996 GSPSAGITQQQLQQSSPQOM 1015

RESULT 5

AI2053

competence protein [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C;Accession: AI2053

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, DNA Res. 8, 2005-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AI2053

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-798 <KUR>

A;Cross-references: UNIPROT:Q9YV14; GB:BA000019; PIDN:BA073682.1; PID:g17131073; GSPDB

A;Experimental source: strain PCC 7120

C;Genetics:

A:Gene: comE

Query Match 6.8%; Score 89; DB 2; Length 798;
Best Local Similarity 21.6%; Pred. No. 5.4;
Matches 44; Conservative 33; Mismatches 73; Indels 54; Gaps 9;

QY 40 EGVSVSVCNTVMSCNINAFSHV-----NIKLRHGQSAIFNEVAPGYFGRDQWQ 92
DB 606 QGVNQIDWAIATDFORNNNDWLEVLQRLAIKNFYAYATNKENSADQAIPQI----- 658
QY 93 LQVQGVSAQLVTKGARDSHAGLYMHVGHQRNNR-----QVTLVSGAEPQSA-----DT 144
DB 659 LQKQGIIVQLLPGV-QTINLGTVAQLINEQPMQLQMLQGLWLVGDVEPKVEVERIMKA 717
QY 145 GFWPVPVAVTAVFILLVALVMAFYRCRCQORREKFFLEPQMKVA---ALRAGAQQ 201
DB 718 GCMPSQVPL-----W-----CNASLKDLYMMLKPOVALIASSGSLSTVLS 759
QY 202 LSRASAEI-----WTPDSE 215
DB 760 LSKTSTKVFVTAQDGAIQWMPNGE 783

RESULT 6
A44048
genome polyprotein - Vilyuisk virus (strain V-1) (fragment)
C:Species: Vilyuisk virus
C:Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 07-Feb-1997
C:Accession: A44048
R:Pritchard, A.E.; Strom, T.; Lipton, H.L.
Virology 191, 469-472, 1992
A:Title: Nucleotide sequence identifies Vilyuisk virus as a divergent Theiler's virus.
A:Reference number: A44048; MUID:93033144; PMID:1413519
A:Accession: A44048
A:Molecule type: genomic RNA
A:Residues: 1-929 <PRI>
A:CROSS-references: GB:M94868
A:Superfamily: foot-and-mouth disease virus genome polyprotein
C:Keywords: polyprotein

Query Match 6.6%; Score 86.5; DB 2; Length 929;
Best Local Similarity 27.0%; Pred. No. 11;
Matches 27; Conservative 16; Mismatches 34; Indels 23; Gaps 5;

QY 61 FSHVNIKL--RAHGQSAIFN-EVAPGYFSRDCWQLOVQGGVAQLVIRKGARDSHAGLYM 117
DB 224 FSHIRVPLPHALAGEHGGVFGATLRHYLAKCGRVQVQCNASQF-----HAGSLIV 275
QY 118 HL-----VGHQRNNRQVTLVSGAEPQSAPTDGF 146
DB 276 FLAPEFYTGTVATSGQBPKNVFLMTTQEPQAP-TGF 314

RESULT 7
S51574
mocr protein - Rhizobium meliloti
C:Species: Rhizobium meliloti
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S51574; S43169
R:Rosbach, S.; Kuipa, D.A.; Rosbach, U.; de Bruijn, F.J.
Mol. Gen. Genet. 245, 11-24, 1994
A:Title: Molecular and genetic characterization of the rhizopine catabolism (mocABRC) gene
A:Reference number: S51569; MUID:95147842; PMID:7845353
A:Accession: S51574
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-493 <ROS>
A:CROSS-references: UNIPROT:P49309; EMBL:X78503; NID:G468758; PID:G468764
C:Genetics:
A:Gene: mocr
C:Superfamily: hypothetical protein b1439

Query Match 6.6%; Score 86; DB 1; Length 493;

Best Local Similarity 26.0%; Pred. No. 5.8;
Matches 53; Conservative 20; Mismatches 65; Indels 66; Gaps 11;

QY 60 AFSHVNIKLRAHGQSAIFNEVA--PGY-----FSRDGWLQVQGGVAQLVIRKGARDSHA 112
DB 96 SLSSRGMAAQPDRRTIPDRIAFHPGYPETKAPFFSTW-----AKLLKRRHARYSHE 147
QY 113 GLYMHVLY-GHQRNNRQVTLVSGAEPQSAPTDGFWPVAVTAVFILLVALVMAFA----- 167
DB 148 DLYGVHWTGHPR-----LKAATAEYLRA-SRGVECAPEQIVVNGTQAALDILARMLV 200
QY 168 -----WYRCRCQORREKFFLEPQMKVAALRAGAQQGLSRASAEI-----WTP 212
DB 201 DEGDCW-----MEEPGY-----IGAQNLSLSAGAKLVPLPVERDGNLS 239
QY 213 DSEPTPRPLALVFKPS-----PLGAL 233
DB 240 EDETRPSRLIFVTPSCQWPLGCL 263

RESULT 8
A82699
conserved hypothetical protein XF1304 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: A82699
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: A82699
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-420 <SIM>
A:CROSS-references: UNIPROT:Q9PDS5; GB:AE003963; GB:AE003849; NID:99106285; PIDN:AAF841
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Laig
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, I.
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.B.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.P.; Marino, C.L.; Marques, M.V.; Martins, P
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1304

Query Match 6.5%; Score 85.5; DB 2; Length 420;
Best Local Similarity 21.1%; Pred. No. 5.4;
Matches 59; Conservative 36; Mismatches 110; Indels 75; Gaps 11;

QY 10 GHVSQAL-GTLLFLAASLSAONEGWDSPICTEGVVSVSGENTVMSCNINAFSHVNIKL 68
DB 33 GTTAQRLEASIALVSQQLGDCPEPMSNP--TGIIILSPDPAKAIGSSDITRVI----- 83
QY 69 RAHGQSAIFNEVAPGYFSRDCWQLOVQGGVAQLVIRKGARDSHAGLYMHLVGHQRNNRQ 128
DB 84 -----RLAPG--ENDLYKLSVADNVAAQAVASGRMSISQGHATLRRLDRELDERS 130
QY 129 VTLVSGAEPQSAPTDGFWPVP--AVVTAVFI-LVALVMAFYRCRCQORREKFFLL 185
DB 131 KMRVFGPLAAGGVAGLVKLPWLDVATAATVGLLIGLL-----TQVTDHRAATR 180
QY 186 EPQMKVALRAG-----AQQGLSRASAEIWTDPDS 214
DB 181 EASEALAAALAGFVATLVATLIGPINLNTVITIASVVVLLPGMLTNAVNELSSQHWVSGT 240

Query 215 EPTPRFLAVFKPSLIGAL-----ELLSPQLFPYAADP 248
 |::||::||| :|:~::~||
Db 241 ARLAGALTIVVFKLT-VGALIIVSLCKLIGLKPOI-YVAQP 278

RESULT 9

B46114
glycoprotein gp13 precursor - equine herpesvirus 1 (strain Kentucky A)
N;Alternate names: glycoprotein C
C;Species: equine herpesvirus 1
A;Note: host Equus caballus (domestic horse)
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C;Accession: B46114
R;Matsumura, T.; Smith, R.H.; O'Callaghan, D.J.
Virolology 193, 910-923, 1993
A;Title: DNA sequence and transcriptional analyses of the region of the equine herpesvir
A;Reference number: A46114; MUID:93212524; PMID:8384760
A;Accession: B46114
A;Molecule type: DNA
A;Residues: 1-468 <MAT>
A;Cross-references: UNIPROT:P12889; GB:S57839; NID:g298846; PIDN:AAB25944.1; PID:g298848
C;Superfamily: herpesvirus glycoprotein F
C;Keywords: glycoprotein; transmembrane protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-468/Product: glycoprotein gp13 #status predicted <GPT>
F:432-451/Domain: transmembrane #status predicted <TMN>
F:46,57,62,92,100,131,203,208,269/Binding site: carbohydrate (Asn) (covalent) #status ph

Query Match 6.5%; Score 85.5; DB 1; Length 468;
Best Local Similarity 22.5%; Fred.No.6.1;
Matches 42; Conservative 24; Mismatches 68; Indels 53; Gaps 7;

QY 2 OTCPLAFFGHVSQAALGTLLFLAASLSAQNEGWDSF-----ICTGVSVSWSGENTVMS 54
 |::||::||| |::||::||| |::||:~::~||
Db 34 QSTP-ATPHITPNLT-TAHGAGSDNTTWANGTESHSHETTICTKSLISVPYYKSVDMN 92

QY 55 CNISNAFSHVNIKLRAHQESAI FNEVARG----VF-----SRDG 90
 |::||:~::~|| |::||:~::~||
Db 93 CTTSVGVMYSERYLIKNLYINORTPFSGTTPGOEENININATKDQTLLLPSTAKRKXSERG 152

QY 91 WLQOV-----QQGVA-QLVTKGARDSHAGLYMMHLV---GHORNNQV 129
 |::||:~::~|| |::||:~::~||
Db 153 GLGVITPDPLPKROLNLPLEHTTEGGTKPELTIKSVDWRTAGIYWWSLAYNGKTLLVNSTSV 212

QY 130 TLEVSGA 136
 |::||:
Db 213 TVSTYNA 219

RESULT 10

GNNYTP

genome polyprotein - murine poliovirus (strain GDDV1)

N;Contains: probable proteinase (EC 3.4.-.-); protein 1A; protein 1B; protein 1C; protein 1D

C;Species: murine poliovirus, Theiler's encephalomyelitis virus

C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999

C;Accession: A29193

R;Pevar, D.C.; Borkowski, J.; Calenoff, M.; Oh, C.K.; Ostrowski, B.; Lipton, H.L.

Virology 165, 1-12, 1988

A;Title: Insights into Theiler's virus neurovirulence based on a genomic comparison of t

A;Reference number: A29193; PMID:88265847; PMID:2838951

A;Accession: A29193

A:Molecule type: Genomic RNA

A;Residues: 1-2303 <PE>

A;Cross-references: GB:M20562; NID:G335221; PIDN:AAA96329.1; FID:gl256538

C;Superfamily: foot-and-mouth disease virus genome polyprotein

C;Keywords: coat protein; core protein; genome-linked protein; hydrolase; polyprotein; i

F;1-76/Domain: leader peptide #status predicted <LDP>

F;77-147/Product: protein 1A (coat protein VP4) #status predicted <VP4>

F;148-414/Product: protein 1B (coat protein VP2) #status predicted <VP2>

F;415-646/Product: protein 1C (coat protein VP3) #status predicted <VP3>

F;647-922/Product: protein 1D (coat protein VP1) #status predicted <VP1>

F;923-1000/Product: protein 2A (core protein P2-3b) #status predicted <P2A>

```

F;1065-1191/Product: protein 2B (core protein P2-5b) #status predicted <P2b>
F;1192-1517/Product: protein 2C (core protein P2-X) #status predicted <P2c>
F;1518-1605/Product: protein 3A (protein p3-1b) #status predicted <P3A>
F;1606-1625/Product: protein 3B (genome-linked protein vpg) #status predicted <P3B>
F;1626-1842/Product: protein 3C (probable proteinase) #status predicted <P3C>
F;1843-2303/Product: protein 3D (probable RNA-directed RNA polymerase) #status predicted
Query Match 6.5%; Score 85.5; DB 1; Length 2303;
Best Local Similarity 27.5%; Pred. No. 39;
Matches 28; Conservative 16; Mismatches 33; Indels 25; Gaps 6;
QY 60 AFSHVNIKLRAH---QGESAIFN-EVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLY 115
|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
DB 223 AFSHIRIPL-PhVLAGEDGGVFGATLRRLHYLCKTGRVQVCNASQF-----HAGSL 273
|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
QY 116 MWHLV-----GHQNNRQVTLVSGAEPSQAPDTGF 146
: : : : : : : : : : : : : : : :
DB 274 LVFMAPERYTKGTSGTWEPSPDFTMDITWRSPQSAP-TGY 314
: : : : : : : : : : : : : : : :
RESULT 11
S13554
genome polypotein - murine poliovirus
C:Species: murine poliovirus, Theiler's encephalomyelitis virus
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S13554
R:Law, K.M.; Brown, T.D.K.
Nucleic Acids Res. 18, 6707-6708, 1990
A:Title: The complete nucleotide sequence of the GDVII strain of Theiler's murine encephalomyelitis virus
A:Reference number: S13554; MUID:91067481; PMID:2251141
A:Accession: S13554
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-2303 <LAW>
A:Cross-references: UNIPROT:Q88595; EMBL:X56019; NID:G62039; PIDN:CAA39496.1; PID:G62040
C:Superfamily: foot-and-mouth disease virus genome polypotein
Query Match 6.5%; Score 85.5; DB 2; Length 2303;
Best Local Similarity 27.5%; Pred. No. 39;
Matches 28; Conservative 16; Mismatches 33; Indels 25; Gaps 6;
QY 60 AFSHVNIKLRAH---QGESAIFN-EVAPGYFSRDGWLQVGGVAQLVIKGARDSHAGLY 115
|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
DB 223 AFSHIRIPL-PhVLAGEDGGVFGATLRRLHYLCKTGRVQVCNASQF-----HAGSL 273
|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
QY 116 MWHLV-----GHQNNRQVTLVSGAEPSQAPDTGF 146
: : : : : : : : : : : : : : : :
DB 274 LVFMAPERYTKGTSGTWEPSPDFTMDITWRSPQSAP-TGY 314
: : : : : : : : : : : : : : : :

```

RESULT 12

A83608 polyamine transport protein PotI PA0304 [imported] - *Pseudomonas aeruginosa* (strain PA01)
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: A83608
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <STO>
A:Cross-references: UNIPROT:Q9I617; GB:AE004468; GB:AE004091; NID:g9946144; PIDN:AAG0361
A:Experimental source: strain PA01
C:Genetics:
C:Gene: potI; PA0304
C:Superfamily: spermidine/putrescine transport system permease protein potI

Query Match 6.5%; Score 84.5; DB 2; Length 289;

Best Local Similarity 26.2%; Pred. No. 4.3;
Matches 60; Conservative 27; Mismatches 91; Indels 51; Gaps 12;

QY 1 MQTCPLAFPGHVSQALGTLFLFLAASLSAQNEGWDSPICTEGVWVSVGWGENTVMSCNISNA 60
DB 100 MVTAPLVWPEVITLISLLLFVA--MAQLIGWPQ---ERGIWTI-WIAHT-----SFC 146
QY 61 FSHVNIKLRAHGQESAIENFVAGPYFSRDGQW---LQVQGGVAQLVKGARDSHAGLYMW 117
DB 147 SSVYVVVVSARLRLEDLSIEAAMDLCAGKPKWVFLITIPMIAPSLAAGGMSFA-LSLD 205
QY 118 HLYGHQRNNQVILEVSG-----AEPSAPDGTGFVPAVTVAVFILLVAL-VMFAYW 169
DB 206 DLV-----LASFVSGSGSTTLPEVFGVAVRLGVKPEINAVASILLISVLSLTFFAWY 257
QY 170 RCRCSQORREKFKFLELPQMKVAALRAGAQQGLSRASAEIWTDPSEPTP 218
DB 258 FTRQAEER-----RRAIQAMEETATD-WQKGS-PTP 288

RESULT 13
VGBEH
N: glycoprotein gp13 precursor - equine herpesvirus 1
N: Alternate names: glycoprotein gC
C: Species: equine herpesvirus 1
C: Date: host Equus caballus (domestic horse)
C: Note: 30-Jun-1989 #sequence revision 30-Jun-1989 #text_change 09-Jul-2004
C: Accession: A28149; A32980; H36796
R: Allen, G. P.; Coogle, L. D.
J. Virol. 62, 2850-2858, 1988
A: Title: Characterization of an equine herpesvirus type 1 gene encoding a glycoprotein (gC)
A: Reference number: A28149; MUID: 88275055; PMID: 2455821
A: Accession: A28149
A: Molecule type: DNA
A: Residues: 1-468 <ALL>
A: Cross-references: UNIPROT: P12889; GB: M86664; NID: G330791; PIDN: AAB02451.1; PID: G330808
A: Experimental source: strain Kentucky T431
A: Note: the authors translated the codon ACA for residue 43 as Pro
R: Guo, P.; Goeble, S.; Davis, S.; Perkins, M. E.; Languet, B.; Desmettre, P.; Allen, G.; F.
J. Virol. 63, 4189-4198, 1989
A: Title: Expression in recombinant vaccinia virus of the equine herpesvirus 1 gene encoding a glycoprotein (gC)
A: Reference number: A32980; MUID: 89382761; PMID: 2550665
A: Accession: A32980
A: Molecule type: DNA
A: Residues: 1-468 <GUO>
A: Cross-references: GB: M86664; NID: G330791; PIDN: AAB02451.1; PID: G330808
R: Telford, E. A. R.; Watson, M. S.; McBride, K.; Davison, A. J.
submitted to Genbank, March 1992
A: Description: The DNA sequence of equine herpesvirus-1.
A: Reference number: A36805
A: Accession: H36796
A: Molecule type: DNA
A: Residues: 1-468 <TEL>
A: Cross-references: GB: M86664; NID: G330791; PIDN: AAB02451.1; PID: G330808
R: Telford, E. A. R.; Watson, M. S.; McBride, K.; Davison, A. J.
Virology 189, 304-316, 1992
A: Title: The DNA sequence of equine herpesvirus-1.
A: Reference number: A41931; MUID: 92295566; PMID: 1318606
A: Contents: annotation; possible protein-coding frames
A: Note: neither amino acid nor nucleotide sequence is given
C: Genetics:
A: Gene: 16
C: Superfamily: herpesvirus glycoprotein F
C: Keywords: Glycoprotein; transmembrane protein
F: 1-30/Domain: signal sequence #status predicted <SIG>
F: 31-468/Product: glycoprotein gp13 #status predicted <MAT>
F: 46, 57, 62, 92, 100, 131, 203, 208, 269/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.4%; Score 83.5; DB 1; Length 468;
Best Local Similarity 22.5%; Pred. No. 9.3;
Matches 42; Conservative 23; Mismatches 69; Indels 53; Gaps 7;

```

Qy 2 QTCPLAPGHVSQMLGTLLFLAASLSAQNEGWDSF-----ICTEGVVSVSWGENTVMS 54
Db 34 QSTP-APFTHTPNLTAAHGAGSDNTNANGTSTHSHETTICTKSLISVPYKSDVMN 92
Qy 55 CNTSNAPSHVNIKLRAHQESAIENEVAPG---YF-----SRDG 90
Db 93 CTTSVGVNYSYRLIYINQRTPSGTPPGDEENYINHNATKDQTLFLFSTAERKKSRRG 152
Qy 91 WQLQV-----QGVVA-QLVIKGARDSHAGLYMMHLV---GHORNNRV 129
Db 153 GQLGVIPIRLPKROLPLPLHTEGGTFPLTIKISVDWRTAGIYVMSLYAKNGTLVNSTSV 212
Qy 130 TLEVSGA 136
Db 213 TVSTYNA 219

RESULT 14
VGBE18
glycoprotein E - human herpesvirus 1
N/Alternate names: US8
C/Species: human herpesvirus 1
C/Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C/Accession: A03733; A45696
R/McGeoch, D.J.; Dolan, A.; Donald, S.; Rixon, F.J.
J. Mol. Biol. 181, 1-13, 1985
A/Title: Sequence determination and genetic content of the short unique region
A/Reference number: A00656; MUID:85160822; PMID:2984429
A/Accession: A03733
A/Molecule type: DNA
A/Residues: 1-550 <MCG>
A/Cross-references: UNIPROT:P04488; GB:X02138; MID:g59865; PIDN:CAA26062.1; PI
R/Experimental source: strain 17
R/Georgopoulou, U.; Michaelidou, A.; Roizman, B.; Mavromara-Nazos, P.
J. Virol. 67, 3961-3968, 1993
A/Title: Identification of a new transcriptional unit that yields a gene produ
A/Reference number: A45696; MUID:93287213; PMID:8389914
A/Accession: A45696
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 438-550 <GEO>
A/Cross-references: GB:S62895; MID:g386127; PIDN:AAB27080.1; PID:g386128
A/Experimental source: R35
A/Note: sequence extracted from NCBI backbone (NCBIN:133646, NCBI:P:133647)
C/Superfamily: herpesvirus glycoprotein E
C/Keywords: Glycoprotein
F,124,243,501/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.3%; Score 83; DB 1; Length 550;
Best Local Similarity 25.4%; Pred. No. 12;
Matches 53; Conservative 16; Mismatches 62; Indels 78; Gaps 11

Qy 80 EVAPGYFSRDGWLQVQGVVAQLVIKGARDSHAGLYMM-----HLVGH-----QK 124
Db 330 EPVPG-----LAWQAASVNLFRDASPQSGSLGLCVVYVNDHIHAWGHITISTAAQY 381
Qy 125 NNRQV-----TLEVSGAESQSPDPTGFVPVAVTAVFLLVALVMEFAYR 170
Db 382 RNAVVEOPLPQRGADLAEPTHPHVGVAPPHTTHAULRGAVMGAA-LLLSALGUSVWAC 440
Qy 171 CRCSQRRKKFFLELPQMKVAALRAGAQQGLSR-----ASAEI---WTPDSE----- 215
Db 441 MTCWRR-----ANRAVKSRAKGKPTYIRVADSELYADWSSDSEGERDQ 485

Qy 216 -----PPTPLALVFKPSPLGA-LELISP 238
Db 486 VPMIAPPERP-----DSPSTNGSGFEILSP 510

```

RESULT 15
E71086
hypothetical protein PH0954 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii

Sun Mar 20 14:17:43 2005

us-10-080-522-1.rpr

C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C/Accession: E71086
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A>Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A/Reference number: A71000; MUID:98344137; PMID:9679194
A/Accession: E71086
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-4436 <KAW>
A/Cross-references: UNIPROT:O58659; GB:AP000004; NID:G3236131; PIDN:BAA30051.1; PID:d103
A/Experimental source: strain OT3
A/Note: this accession replaces an interim accession for a sequence replaced by GenBank
C/Genetics:
A/Gene: PH0954

Query Match 6.3%; Score 83; DB 2; Length 4436;
Best Local Similarity 23.0%; Pred. No. 1.4e+02;
Matches 37; Conservative 24; Mismatches 58; Indels 42; Gaps 8;

QY 34 DSPICTEGVVS-----WGENTVMSCNISNAFSHVNIKLRA-----HGQE 74
DB 300 DDPDLNNLAEGVWPGDYWENASINNLI PGEFASINFKVRTTSKIPSAKVLLRNGVE 359

QY 75 SAI-----FNEVAPGYFSRDGWLQVGG--VAQLVIKGARDSHA-----GLYWHILV 120
DB 360 EKIEYLSFYNGIAEGEIS--W--LVQGGNYTLALIVEGKGIDINSNNIYLLGNYNFPLP 414

QY 121 GHORNNRQVTL-----EVSGAEPOSAPDTGFWPVPVAVTAVF 157
DB 415 NFEVGNYSIDLPTCVDSTGEVRVNTSTANWSIPVRLTLVY 455

Search completed: March 20, 2005, 09:49:42
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 20, 2005, 06:56:18 ; Search time 92 Seconds
(without alignments)
1380.387 Million cell updates/sec

Title: US-10-080-522-1
Perfect score: 1308
Sequence: 1 MQTCPLAPGHVSQALGTL.....PLGALELLSQPLFPYAADP 248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1308	100.0	248	1 SCTM_HUMAN	Q8wv66 homo sapien
2	382	29.2	192	2 Q921W8	Q921W8 mus musculus
3	376.5	28.8	218	2 Q6P781	Q6P781 rattus norv
4	358	27.4	201	2 Q6AYS0	Q6AYS0 rattus norv
5	327.5	25.0	211	2 Q8YDU7	Q8YDU7 mus musculus
6	326	24.9	212	1 SCTM_MOUSE	Q9J159 mus musculus
7	325	24.8	212	2 Q9D966	Q9D966 mus musculus
8	324	24.8	212	2 Q8CE38	Q8CE38 mus musculus
9	103.5	7.9	550	2 Q52289	Q52289 pseudomonas
10	102.5	7.8	506	1 SHS1_BOVIN	O46631 bos taurus
11	101	7.7	334	2 Q8XZ99	Q8XZ99 ralstonia s
12	98.5	7.5	304	2 Q8BPN5	Q9DPn5 caenorhabdi
13	98	7.5	542	2 Q8NHNS	Q8nhns homo sapien
14	97	7.4	894	2 Q9V5F7	Q9V5f7 drosophila
15	96.5	7.4	550	2 Q8BP09	Q8BP09 pseudomonas
16	96	7.3	442	2 Q8HB33	Q8hb33 homo sapien
17	96	7.3	917	2 Q6NNX3	Q6nnx3 drosophila
18	96	7.3	969	2 Q8KKG6	Q8Kkg6 homo sapien
19	95	7.3	1140	2 Q80T91	Q80t91 mus musculus
20	94.5	7.2	393	2 Q82004	Q82004 lycopersico
21	94.5	7.2	690	2 Q62AS3	Q62as3 burkholderi
22	94.5	7.2	690	2 Q63JU4	Q63ju4 burkholderi
23	91.5	7.0	864	2 Q6P779	Q6P779 rattus norv
24	91.5	7.0	1260	2 Q7NR05	Q7nr05 chromobacte
25	91	7.0	6620	2 Q96AA2	Q96aa2 homo sapien
26	90.5	6.9	271	2 Q9S218	Q9S218 streptomyce
27	90	6.9	1088	2 Q9C9D7	Q9C9d7 arabidopsis
28	89	6.8	798	2 Q8YVJ4	Q8yvJ4 anabaena sp
29	88.5	6.8	352	2 Q63NY8	Q63ny8 burkholderi
30	88	6.7	327	2 Q8RIG9	Q8ryg9 oryza sativ
31	87.5	6.7	298	2 Q804R4	Q804r4 brachydanio

32	87.5	6.7	372	2	Q90Y50	Q90y50 brachydanio
33	87.5	6.7	645	2	Q8MJZ5	Q8mjz5 pan troglod
34	87.5	6.7	947	2	Q8BKK7	Q8bkk7 mus musculus
35	87	6.7	259	2	Q82SL8	Q82sl8 nitrosomona
36	86.5	6.6	395	2	Q88LS8	Q88ls8 pseudomonas
37	86.5	6.6	420	2	Q87DX4	Q87dx4 xyella fae
38	86.5	6.6	930	2	Q02472	Q02472 theiller's e
39	86	6.6	493	1	MOCR_RHIME	P49309 rhizobium m
40	86	6.6	508	2	Q80ZD5	Q80zd5 rattus norv
41	85.5	6.5	407	2	Q9D2J4	Q9d2j4 mus musculus
42	85.5	6.5	418	2	Q8MI84	Q8mi84 pongo pygma
43	85.5	6.5	420	2	Q9PDS5	Q9pd5 xyella fae
44	85.5	6.5	430	2	Q8MI85	Q8mi85 pongo pygma
45	85.5	6.5	922	2	Q88495	Q88495 theiller's e

ALIGNMENTS

RESULT 1
SCTM_HUMAN
ID SCTM_HUMAN STANDARD; PRT; 248 AA.
AC Q8WV6; O00466;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Secreted and transmembrane protein 1 precursor (Protein K12).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=98149980; PubMed=9480746; DOI=10.1006/geno.1997.5151;
RT Sientz-Kessler K.A., Hale L.P., Kaufman R.E.;
RT "Identification and characterization of K12 (SCTM1), a novel human
RT gene that encodes a Golgi-associated protein with transmembrane and
RT secreted isoforms.";
RN Genomics 47:327-340(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP INTERACTION WITH CD7.
RX MEDLINE=20119303; PubMed=10652336; DOI=10.1074/jbc.275.5.3431;
RA Lyman S.D., Escobar S., Rousseau A.-M., Armstrong A., Fanslow W.C.;
RT "Identification of CD7 as a cognate of the human K12 (SCTM1)
RT protein.";
RL J. Biol. Chem. 275:3431-3437(2000).
CC -!- SUBUNIT: Interacts with CD7.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable). Also

[illegible]

RESULT	3
Q6P781	
ID	Q6P781 PRELIMINARY; PRT; 218 AA.
AC	Q6P781;
DT	05-JUL-2004 (TREMBLrel. 27, Created)
DT	05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE	Hypothetical protein MGC72571.
GN	Name=MGC72571;
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX	NCBI_TaxID=10116;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Prostate;
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altshul S.F., Zeeberg B., Buetow K.H., Schafer C.P., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

01-NOV-2002 (Zemaburel. 25, last sequence update)
01-OCT-2003 (ZEMABUREL. 25, last annotation update)

DT NCBI_TaxID=10090;
DE Sectm1 protein.
GN Name=Seetm1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Mammary tumor;
RX MEDLINE=22389257; Pubmed=12477933; DOI=10.1073/pnas.242603899;
RA Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.C., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Iqbalnaro N.A., Peters G.J., Abramson R.D., Mollany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,

01-NOV-2002 (Zemaburel. 25, last sequence update)
01-OCT-2003 (ZEMABUREL. 25, last annotation update)

DT NCBI_TaxID=10090;
DE Sectm1 protein.
GN Name=Seetm1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Mammary tumor;
RX MEDLINE=22389257; Pubmed=12477933; DOI=10.1073/pnas.242603899;
RA Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.C., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Iqbalnaro N.A., Peters G.J., Abramson R.D., Mollany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,

Jones S.J., Marra M.A.;
 "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CZECH II; TISSUE=Mammary tumor;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBAJ databases.
 DR EMBL; BC020159; AAH20159.1; -;
 DR MGD; MGI:1929083; Sectm1.
 DR SMART; SM00409; IG; 1.
 SQ SEQUENCE 211 AA; 23433 MW; 9E6FB436592B2D8 CRC64;
 Query Match 25.0%; Score 327.5; DB 2; Length 211;
 Best Local Similarity 40.0%; Pred. No. 6.9e-21;
 Matches 78; Conservative 24; Mismatches 70; Indels 23; Gaps 3;
 QY 10 GHVSQALGTLFLAASLSAQNEDGWDSPICTEGVSVSGENTVMSCNISNAPSHVNIKL 69
 DB 10 GLFPRMLWALLLLAASLNAYNHVWDKPCCTEHEVSVNRGSRVYVMAACNINLNRDVTIELV 69
 QY 70 AHQGESALFNEVAPGYFSDRGQVQGVQVQLVKGARDSHAGLYMHLVGHQNNRQV 129
 DB 70 TSEKTSIIIFNTPPGNYSKDSWQLHIQGVQVQLVITDAQCKSGEYWKRLGLQAEFKNF 129
 QY 130 TLVSGAEPQAPDPTGFWPVP-----AVVTAVFILLVA-----LVMPFATWRCRCQQ 177
 DB 130 NLIVNAADQKTEDELPVTKVPDPTAVMTVEVIIIAITATITITIGVFWVY----- 181
 QY 178 REKKFFLEPQMKVA 192
 DB 182 ---KQFPVAPQIQMS 193
 RESULT 6
 SCTM MOUSE STANDARD; PRT; 212 AA.
 ID SCTM MOUSE
 AC Q9JL59;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Secreted and transmembrane protein 1 precursor.
 GN Name=Secretm;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND INTERACTION WITH CD7.
 RC TISSUE=Colon;
 RX MEDLINE=201119303; PubMed=10652336; DOI=10.1074/jbc.275.5.3431;
 RA Lyman S.D., Escobar S., Rousseau A.-M., Armstrong A., Fanslow W.C.;
 RT "Identification of CD7 as a cognate of the human K12 (SECTM1)
 protein";
 RL J. Biol. Chem. 275:3431-3437 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk R.W.,
 RA Villalón D.K., Muny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RL -I- SUBUNIT: Interacts with CD7.
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein (Probable). Also
 CC found as secreted (By similarity).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF210700; AAF30406.1; -;
 DR EMBL; BC010805; AAH10805.1; -;
 DR MGD; MGI:1929083; Sectm1.
 DR InterPro; IPR003599; IG-like.
 DR SMART; SM00409; IG; 1.
 KW Signal; Transmembrane.
 FT SIGNAL 1 28
 FT CHAIN 29 212
 FT DOMAIN 29 160
 FT TRANSMEM 161 181
 FT DOMAIN 182 212
 FT DISULFID 38 55
 FT CARBOHYD 56 56
 FT CARBOHYD 85 85
 FT CARBOHYD 114 114
 FT CARBOHYD 130 130
 FT SEQUENCE 212 AA; 23477 MW; 75113E877A2C5B87 CRC64;
 SQ
 Query Match 24.9%; Score 326; DB 1; Length 212;
 Best Local Similarity 37.8%; Pred. No. 9.4e-21;
 Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;
 QY 10 GHVSQALGTLFLAASLSAQNEDGWDSPICTEGVSVSGENTVMSCNISNAPSHVNIKL 69
 DB 10 GLFPRMLWALLLLAASLNAYNHVWDKPCCTEHEVSVNRGSRVYVMAACNINLNRDVTIELV 69
 QY 70 AHQGESALFNEVAPGYFSDRGQVQGVQVQLVKGARDSHAGLYMHLVGHQNNRQV 129
 DB 70 TSEKTSIIIFNTPPGNYSKDSWQLHIQGVQVQLVITDAQCKSGEYWKRLGLQAEFKNF 129
 QY 130 TLVSGAEPQAPDPTGFWPVP-----AVVTAVFILLVA-----LVMPFATWRCRCQQ 176
 DB 130 NLIVNAADQKTEDELPVTKVPDPTAVMTVEVIIIAITATITITIGVFWVY----- 182
 QY 177 REKKFFLEPQMKVA 192
 DB 183 ---KQFPVAPQIQMS 194
 RESULT 7
 Q9D966 PRELIMINARY; PRT; 212 AA.
 ID Q9D966
 AC Q9D966; (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TEMBLrel. 26, Last annotation update)
 DE Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched
 DE library, clone:1810003C24 product:secreted and transmembrane 1, full
 DE insert sequence.
 GN Name=Secretm;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muranatsu M., Inoue Y., Kira A., Hayashizaki Y.; "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer."; Genome Res. 10:1757-1771(2000).

[6]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Skin;

RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., Saito K., Saitoh K., Shingawa A., Shiraki T., Sogabe Y., Tagami M., Sasaki D., Saibata K., Takaku-Akahira S., Takeda Y., Tanaka T., Tagawa A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.; Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RL ENBL; AK029082; BAC26284.1; --

DR MGD; MGI:1929083; Sectml.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR InterPro; IPR003599; IG.

DR InterPro; IPR007110; IG-like.

DR SMART; SM00409; IG; 1.

KW Transmembrane.

SQ SEQUENCE 212 AA; 23559 MW; C863064DCE36EB7E CRC64;

Query Match 24.8%; Score 324; DB 2; Length 212;

Best Local Similarity 37.8%; Pred. No. 1.4e-20;

Matches 74; Conservative 25; Mismatches 73; Indels 24; Gaps 2;

QY 10 GHVSQALGTLFLAASLSAQNQEGWDSPICTEGVSVSWGNTVMSNINAFSHVNIKL 69

DB 10 GLFPRMLWALLLLAASLNAYNHVWDKPCCTEHEVSVNRGSRVVMACNINLNLDVTIELV 69

QY 70 AHGQESAIENFVAPGYFSDQWLOVQGVVAOLVIKGRDASHAGLVNHLVGHQRNROV 129

DB 70 TSKTSTIIFNKTTPNGYSKDSWQLHIQGGQALVITDAQKHGGEVWKLGRGFQAEKFN 129

QY 130 TLEVSGAEPOSAPDTGFVFPV-----AVTVAFILLVALVMPFAYRCRCQQ 176

DB 130 NLIVNAADRQKTELDLVTKVDPKPTAVRTEVIIIIAIIITIGIGVEVWY----- 182

QY 177 RREKKFFLLPEQMKVA 192

DB 183 ----KQFPVAFQIQMS 194

RESULT 9

Q52289 PRELIMINARY; PRT; 550 AA.

AC Q52289

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE UxpA protein.

GN Name=uxpA;

OS Pseudomonas putida.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

OX NCBI_TaxID=303;

RN [1]

RN SEQUENCE FROM N.A.

RC STRAIN=WCS358;

RA MedLine=96186881; PubMed=8602167; DOI=10.1007/s004380050102;

RX de Groot A., Kriliger J.J., Filloux A., Tommassen J.; "Characterization of type II protein secretion (xcp) genes in the plant growth-stimulating Pseudomonas putida, strain WCS358."; Mol. Gen. Genet. 250:491-504(1996).

RN [2]

RN SEQUENCE FROM N.A.

RC STRAIN=WCS358;

RA de Groot A.; Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

RL ENBL; X81085; CAA56977.1; --

DR PIR; S64725; S64725.

SQ SEQUENCE 550 AA; 59118 MW; 684D79D0D279868B CRC64;

Query Match 7.9%; Score 103.5; DB 2; Length 550;

Best Local Similarity 23.2%; Pred. No. 1.4;

Matches 72; Conservative 42; Mismatches 118; Indels 79; Gaps 16;

QY 2 QTCPLAPPGHVSOA-IGTLFLAA-----SLSAQN-EGWDSPICTEGVSVS 46

DB 90 QLAPLLDAGRADEANLGGYAVLGLLEQLRGEAGAGNSLTLENGQGWG-----SGLAYLT 145

QY 47 WGENVTMSCNISNFAFHVNIKLRA--HGQESAIENFVAPGYFSR-----DGWQ 92

DB 146 QGSGVQSGQLLSEARVSDERVLFPQRSAALYRQASATTLGAGLADQRQALGLEPLQ 205

QY 93 LOVGQGVAAQLVIKQ-----ARDSHAGLVMMH-----LVGHQRNNRVTLVSGAEPOSAP-- 142

DB 206 LFERGG-ARIATVGVTDVPAQDQKASLKQWYQALLPVFQQRREADLVVALADVGTGPGL 264

QY 143 -----DTGFWPVPVAVTVFILLVALVMEFAM-----VRCRC-----SQQ 176

DB 265 WLAERLPAIDLLLCARGODLWPTVQATQASGRVP-VLPAGCGSGAFELRCQVAGOW 323

QY 177 RREKKFF-----LLEPQMKVAAALFAGAQGLSRASAELEWTPDSEPTPR-PLALVFKPSPL 230

DB 324 QFEGRPFTFETLSPAAQVRAQGLQALNQQRAGHAWL--DQPLARAPQALWRRDTRG 381

231 GALELLSPQL 241

382 GSWDRLLHQAL 392

RESULT 10

SHS1_BOVIN

ID SHS1_BOVIN STANDARD; PRT; 506 AA.

AC Q46631; Q46632;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Protein-tyrosine phosphatase non-receptor type substrate 1 precursor (SHP substrate-1) (SHPs-1) (Inhibitory receptor SHPS-1) (Signal-regulatory protein alpha-1) (SIRP-alpha-1) (MyD-1 antigen).

DE Name=PTPNS1; Synonyms=MYD1, SHPS1, SIRP;

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RN SEQUENCE FROM N.A., AND VARIANTS SER-23; ALA-28; LEU-61; ARG-70;

RP HIS-120; 125-GLN; 129-GLY; 129-HIS; 132-VAL; ASN-145; VAL-153;

RP ASP-203; ARG-261; LEU-302; LEU-316; ARG-337; ASN-367; LEU-422; PHE-429

AND GLU-433.

RC STRAIN=Friesian; TISSUE=Peripheral blood;

RX MEDLINE=98143722; PubMed=9485180;

RX DOI=10.1002/(SICI)1521-4141(199801)28:01<1::AID-IMMU1>3.3.CO;2-M;

RA Brooke G.P., Parsons K.R., Howard C.J.; "Cloning of two members of the SIRP alpha family of protein tyrosine phosphatase binding proteins in cattle that are expressed on monocytes and a subpopulation of dendritic cells and which mediate binding to CD4 T cells.";

RT Eur. J. Immunol. 28:1-11(1998).

CC -!- FUNCTION: Immunoglobulin-like cell surface receptor for CD47. Acts as docking protein and induces translocation of PTPN6, PTPN11 and other binding partners from the cytosol to the plasma membrane. Supports adhesion of cerebellar neurons, neurite outgrowth and glial cell attachment. May play a key role in intracellular signaling during synaptogenesis and in synaptic function. Involved in the negative regulation of receptor tyrosine kinase-coupled

cellular responses induced by cell adhesion, growth factors or insulin. Mediates negative regulation of phagocytosis, mast cell activation and dendritic cell activation. CD47 binding prevents maturation of immature dendritic cells and inhibits cytokine production by mature dendritic cells (By similarity).

-!- SUBUNIT: Binds PRPN1 when tyrosine-phosphorylated, except in macrophages, where it primarily binds PRPN6. Binds GRB2 in vitro. Binds JAK2 irrespective of its phosphorylation status and forms a stable complex. Binds SCAP1 and/or SCAP2. The resulting complex recruits FYB. Binds FGR and PTK2B (By similarity).

-!- SUBCELLULAR LOCATION: Type I membrane protein.

-!- TISSUE SPECIFICITY: Highly expressed in spleen macrophages. Detected in skin dendritic cells.

-!- PTM: Phosphorylated on tyrosine residues (By similarity).

-!- SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.

-!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Y11045; CAA71942.1; -.
 EMBL; Y11046; CAA71943.1; -.
 HSP; O88792; 1F97.
 InterPro; IPR007110; Ig-like.
 InterPro; IPR003597; Ig_C1.
 InterPro; IPR003006; Ig_MHC.
 Pfam; PF00047; Ig; 3.
 SMART; SM00407; IG_C1; 2.
 PROSITE; PS50835; IG LIKE; 3.
 PROSITE; PS00290; IG MHC; FALSE NEG.
 Glycoprotein; Immunoglobulin domain; Phosphorylation; Polymorphism;
 Repeat; SH3-binding; Signal; Transmembrane.
 FT SIGNAL 1 29
 FT CHAIN 30 506
 FT DOMAIN 30 371
 FT TRANSMEM 372 392
 FT DOMAIN 393 506
 FT DOMAIN 30 145
 FT DOMAIN 148 248
 FT DOMAIN 255 348
 FT DISULFID 55 121
 FT DISULFID 170 228
 FT SITE 273 331
 FT SITE 432 435
 FT SITE 441 446
 FT SITE 455 458
 FT SITE 472 475
 FT SITE 498 501
 FT MOD_RES 431 431
 FT MOD_RES 455 455
 FT MOD_RES 472 472
 FT MOD_RES 498 498
 FT CARBOHYD 92 92
 FT CARBOHYD 167 167
 FT CARBOHYD 179 179
 FT CARBOHYD 204 204
 FT CARBOHYD 210 210
 FT CARBOHYD 246 246
 FT CARBOHYD 270 270
 FT CARBOHYD 292 292
 FT CARBOHYD 311 311
 FT CARBOHYD 319 319
 FT CARBOHYD 344 344

FT CARBOHYD 365 365 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 368 368 N-linked (GlcNAc...) (Potential).
 FT VARIANT 23 23 N -> S.
 FT VARIANT 28 28 T -> A.
 FT VARIANT 61 61 S -> L.
 FT VARIANT 70 70 K -> R.
 FT VARIANT 120 120 Y -> H.
 FT VARIANT 125 125 R -> Q.
 FT VARIANT 127 127 E -> G.
 FT VARIANT 129 129 M -> H.
 FT VARIANT 132 132 M -> V.
 FT VARIANT 145 145 S -> N.
 FT VARIANT 153 153 L -> V.
 FT VARIANT 203 203 G -> D.
 FT VARIANT 261 261 N -> R.
 FT VARIANT 302 302 F -> L.
 FT VARIANT 316 316 F -> L.
 FT VARIANT 337 337 G -> R.
 FT VARIANT 367 367 S -> N.
 FT VARIANT 422 422 Q -> L.
 FT VARIANT 429 429 I -> F.
 FT VARIANT 433 433 D -> E.
 SQ SEQUENCE 506 AA; 55093 MW; 6B7E310677FCE9CB CRC64;

Query Match 7.8%; Score 102.5; DB 1; Length 506;
 Best Local Similarity 27.0%; Pred. NO. 1.5;
 Matches 51; Conservative 25; Mismatches 82; Indels 31; Gaps 7;

QY 4 CPLAPPCHVSQALCTLLFLAASLSAQNEGWSPTCTEGVSVSWGENTVMSCNISNAP-S 62
 DB 228 CEVA---HVTIQGGPPLRGNTANLS---ETIRVPPTLEITGSPSAGNVNVTQVKNKFPYR 281
 QY 63 HVNIKLRAGH-----QESAIFNEVAPGYFSRDGMOLQVGGVAQLVIKGRDASHAGLYMW 117
 DB 282 HLQLTWLENGNMSRTEAASVFVENKDTGTFNQTSLVFNSSAHEAVLTCQVEHD----- 336
 QY 118 HLVGQRNRRQVTLVSGAEPQAPDTGTFWVPV-----AVTAVPILILVALVMFAW 168
 DB 337 ---GQPAVSKNHTLEVSA--PQODDTGTGTPGNDNNNTSIFIVGVVVCALLVALLIAL 391
 QY 169 YRCRCSQOR 177
 DB 392 YLLRIRQNK 400

RESULT 11
 Q8XZ99 PRELIMINARY; PRT; 334 AA.
 AC Q8XZ99
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE PROBABLE ALCOHOL DEHYDROGENASE-LIKE OXIDOREDUCTASE PROTEIN (EC 1.1.1.-)
 DE)
 GN Name=RS03800; OrderedLocusNames=RS031505;
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM11000;
 RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brottier P., Camus J.-C., Cattolico L.,
 RA Chander M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schief T.,
 RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 FT "Genome sequence of the plant pathogen Ralstonia solanacearum";
 RL Nature 415:497-502(2002).
 CC -!- COPACTOR: Zinc (By similarity).
 CC -!- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase

CC family.
DR EMBL: AL646064; CAD15207.1; --
DR HSP; P75691; 1UP
DR GO: GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0008270; F:zinc ion binding; IEA.
DR Pfam: PF00107; ADH zinc N; 1.
DR PROSITE; PS00059; ADH ZINC; UNKNOWN 1.
DR Complete proteome; Meta1-binding; Oxidoreductase; Zinc.
KW SEQUENCE 334 AA; 35204 MW; 0F0176E32E8F3BA2 CRC64;
SQ
Query Match 7.7%; Score 101; DB 2; Length 334;
Best Local Similarity 24.6%; Pred. No. 1.3;
Matches 67; Conservative 28; Mismatches 83; Indels 94; Gaps 16;
QY 5 PLAFCHVSQALGTLFLAASLSAQNEGWDSPTCTEGVSVSWGENTVMSNCISNAFASHV 64
DB 54 PALIPGH--EIVGTVDACGPGVSLTGE-----RVGVPWLGSCTGCTPCF-- 98
QY 65 NIKLRAHGOESAIENFVAFYFSRDGWLQVQGGVAQLVIKAR-----DSHAG- 113
DB 99 -----RDQENLCDRPFTGY-TRD-----GGVAEYTVCDARYCLPIPARYDDAHAAP 144
QY 114 LYMHVLVGHQNRNRQVTLVSGAEPOSAPDTGFWPVPVAVTAVFILLVALVMFAWYRCRC 173
DB 145 LLCAGLIGYR-----TLRMAG-DARRLGIYFGAAHLVTQIAV----- 182
QY 174 SQORREKFPLEPPOMKVALRAGAACQGLSRASAEIWTDPSE-PTPRPL--ALVFKPSP 230
DB 183 AEQREVAF-----TRAGDTGAQO--LAROTGACWAGASEAPPAPLDAALIF--APV 231
QY 231 GALELLSPQ-----PLFFY 244
DB 232 GALVPLALQAVKVGTVCGGIHMSDIPAPPY 263
RESULT 12
Q9BPN5 PRELIMINARY; PRT; 304 AA.
AC Q9BPN5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Y32G9A.8.
GN Name=Y32G9A.8; ORFNames=Y32G9A.8;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Bemis G., Courtney L.;
RT "The sequence of C. elegans cosmid Y32G9A.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC087794; AAG53701.2; --
DR WormBase; WEGene00021305; Y32G9A.8.
DR WormPep; Y32G9A.8; CE31101.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Hypothetical protein_IG_LIKE; 1.
KW SEQUENCE 304 AA; 34112 MW; AEF48BB42A0DB012 CRC64;
SQ
Query Match 7.5%; Score 98.5; DB 2; Length 304;
Best Local Similarity 22.1%; Pred. No. 1.9;
Matches 36; Conservative 28; Mismatches 52; Indels 47; Gaps 6;
QY 49 ENTVMSCNISNAESHVNIK-----LRAHGOESAIENFVAFYFS----- 87
DB 121 DNSIIIGNMF-AYSHVPVKKNNETWELKSESEFTVGVPAVAPLDSMAARIQCPVGYPE 179
QY 88 -----RDGWLQVQGGV-----AQLVIKARGDSHAGLYMHVLVGHQNRNRQVTLVSGAE 137
DB 180 PQIWIYKDKFLEIGRVKFTAGVLSIEGAQEDAGVYRCEA-----TNQFFVQIDGPE 233
QY 138 PQSA-----PDTGFWPVPVAVTAVFILLVALVMFAWYRC 171
DB 234 QHFAVKLDQELRIGDSYGMPLAILIILLFLVFTQCRC 276
RESULT 13
Q8NHNS PRELIMINARY; PRT; 542 AA.
ID Q8NHNS;
AC Q8NHNS;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Obscurin (fragment).
GN Name=OBSCN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Young P., Ehler E., Gautel M.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Gautel M.S.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ314903; CAC85750.1; --
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS50835; IG_LIKE; 6.
FT NON_TER 1 542
FT NON_TER 542 542
SQ SEQUENCE 542 AA; 58725 MW; F2F41C47CCF4F198 CRC64;
Query Match 7.5%; Score 98; DB 2; Length 542;
Best Local Similarity 23.2%; Pred. No. 4.1;
Matches 60; Conservative 35; Mismatches 94; Indels 70; Gaps 11;
QY 11 HVSQALGTLFLAASLSAQNEGWDSPTCTEGV-----VSVWGENTVMSNCISNA 60
DB 248 HGAQVLDIAIY---SCRVGAEGQDPFVQVEEAAKFCRLLEPVCGLGTGTVLACELSPA 304
QY 61 FSHV-----NIKLRHGOESAIENFVAFYFSRDGWLQVQGGVAQLVIKARGDSHAGLY 115
DB 305 CAEVMVRCGNTQLRV-GKR-----FQMAEGFVRSITVLGLRAEDAGEY 347
QY 116 MMHLVGHQNRNR---QVTLV-----SCAEPQSPADTGFVFPVAVTAVFILLVALVM 165

Db 348 ----VCESRDDHTSQAQLTVSVPRVVKFMSGLSTVVAEGEGENTFQCVSPSDVAVV---- 399

Qy 166 FAWYRCRCQORREKKFFLLPQMKVAALRAGAQQGLSRASAE-----WTPDSEPTPRP 220

Db 400 --WFRDGA-----LLQPSKFAISQSGASHSLTISDLVLEDAQGITVEAGSS 447

Qy 221 LALVFKPSPGLALELLSPQ 239

Db 448 AALRVREAPVLFKKLEPQ 466

RESULT 14

Q9V5F7

ID Q9V5F7 PRELIMINARY; PRT; 894 AA.

AC Q9V5F7;

DT 01-MAY-2000 (TRENBLrel. 13, Created)

DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE CG2292-PA.

GN ORFNames=CG2292;

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,

RA April J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C.J., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskae R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,

RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,

RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou S., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.;"

RL Science 287:2185-2195(2000).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=22426065; PubMed=12537568;

RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,

RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,

RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,

RA Paciele J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,

RA Svirskae R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,

RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;

RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila

RT melanogaster euchromatic genome sequence.;"

RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=22426070; PubMed=12537573;

RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskae R.,

RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,

RA Ashburner M., Celniker S.E.;

RT "The transposable elements of the Drosophila melanogaster euchromatin:

RT a genomics perspective.;"

RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=22426069; PubMed=12537572;

RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,

RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,

RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,

RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,

RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,

RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,

RA Lewis S.E.;

RT "Annotation of the Drosophila melanogaster euchromatic genome: a

RT systematic review.;"

RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).

RN [5]

RP SEQUENCE FROM N.A.

RG FlyBase;

RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

RN [6]

RP SEQUENCE FROM N.A.

RG FlyBase;

RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE003831; AAF58853.2; --

DR FlyBase; FEGN0033479; CG2292.

DR InterPro; IPR007070; PIGN.

DR Pfam; PF04987; PIGN; 1.

SQ SEQUENCE 894 AA; 101828 MW; 9BD68206E9595187 CRC64;

Query Match 7.4%; Score 97; DB 2; Length 894;

Best Local Similarity 22.7%; Pred. No. 8.8;

Matches 62; Conservative 33; Mismatches 88; Indels 90; Gaps 15;

Qy 9 PCHVSOALGTLFLAASLSAQNEGWS-PICTEGVVSVS-----WGENTVMSCNISNAPS 62

Db 95 PGHITLIG-----LYEDPSAVLRGWSNPIDPDTVFNRSQTYANGANDVL-----NVFS 145

Qy 63 HVN-----IKLRAHQESAIFFNEVAPGY--FSRDGWLQVQGVQVQLVKGARDSHAGLYM 116

Db 146 HVSNGGEINLRFYNHDL---DFSPGYDAYEEDWVFKVKLLQKREALQRAQNVVVF 201

Qy 117 WHLVGHQRNQVTLVSVG-REPOSAPDTGFVPVAVTAVFILLVALVMPFAYWRCRSQ 175

Db 202 LHLIG-----LDTAGHVHKFGAP-----KFRRTL 225

Qy 176 QRREK-----KPFLEPQMKVAALRAGAQQGLSRASAE-----LW-TPDS 214

Db 226 EKTEKGVVAIQEFERVPDKRTAYLLT-ADHGMTDSGAHSGSGSPHETDTPPMLNGAGAS 284

Qy 215 EPTPRPLALVFKPS-----PLGALELSPOPL 241

Db 285 RAVPKPGGRTMPNNEGPAFMLEQAQLTPL 317

RESULT 15

Q88P09

ID Q88P09 PRELIMINARY; PRT; 550 AA.

AC Q88P09;

DT 01-JUN-2003 (TRENBLrel. 24, Created)

DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)

Search completed: March 20, 2005, 09:49:13
Job time : 100 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 00:33:52 ; Search time 9530 Seconds
(without alignments)
11084.194 Million cell updates/sec

Title: US-10-080-522-2
Perfect score: 2180
Sequence: 1 ATTCTGCTTCTTTAGCGT.....TTAAATAAGTTTCGTCGC 2180

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEnbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2180	100.0	2180	6	AR071760 Sequence
2	2180	100.0	2180	6	AR097657 Sequence
3	2180	100.0	2180	6	AR195192 Sequence
4	1951	89.5	2000	6	CO776411 Sequence
5	1951	89.5	2000	6	CO861464 Sequence
6	1951	89.5	2000	6	AX055560 Sequence
7	1951	89.5	2000	6	AX333755 Sequence
8	1951	89.5	2000	9	HSU77643 Homo sapien
9	1681.2	77.1	1761	9	BC017716 Homo sapien
10	1230.6	56.4	154252	9	AC132938 Homo sapi
11	1230.6	56.4	171569	9	AC132872 Homo sapi
12	938.6	43.1	1464	6	BD270355 50 human
13	912.8	41.9	1064	6	BD130223 Human sig
14	732.8	33.6	747	12	BF008191 Synthetic
15	732.6	33.6	166458	2	AC084737 Homo sapi
c 16	564.8	25.9	66324	2	AC134680 Homo sapi
17	506.2	23.2	592	6	AX780236 Sequence
c 18	462.4	21.2	71352	2	AC137090 Homo sapi
19	363.8	16.7	401	6	AX778685 Sequence

20	195.4	9.0	1629	10	BC078937	BC078937 Rattus no
21	189	8.7	1582	10	BC010462	BC010462 Mus muscu
22	179	8.2	229839	2	AC094588	AC094588 Rattus no
c 23	179	8.2	249664	2	AC091384	AC091384 Rattus no
c 24	179	8.2	260404	2	AC128909	AC128909 Rattus no
25	174.6	8.0	1249	10	BC061794	BC061794 Rattus no
c 26	170.8	7.8	198011	10	AL662887	AL662887 Mouse DNA
27	159.6	7.3	1182	10	BC020159	BC020159 Mus muscu
28	157	7.2	1076	6	CO777327	CO777327 Sequence
29	156.8	7.2	71352	2	AC137090	AC137090 Homo sapi
30	155.4	7.1	1112	6	CO777326	CO777326 Sequence
31	155.4	7.1	1112	6	AX055564	AX055564 Sequence
32	155.4	7.1	1112	10	AF210700	AF210700 Mus muscu
33	155.4	7.1	1273	10	BC010805	BC010805 Mus muscu
c 34	139	6.4	73178	2	AC134665	AC134665 Homo sapi
35	77	3.5	125020	9	AF429315	AF429315 Homo sapi
c 36	72.8	3.3	125020	9	AF429315	AF429315 Homo sapi
37	62.4	2.9	101509	2	AC027353	AC027353 Homo sapi
38	61.6	2.8	283769	2	AC094777	AC094777 Rattus no
39	60.8	2.8	82815	2	AC062001	AC062001 Homo sapi
c 40	60.6	2.8	250046	2	AC113592	AC113592 Mus muscu
41	60.4	2.8	300695	2	AC079431	AC079431 Mus muscu
42	60	2.8	60	6	CO541600	CO541600 Sequence
43	59.4	2.7	53552	2	AC100796	AC100796 Homo sapi
44	58.8	2.7	222466	2	AC136508	AC136508 Pan trogl
c 45	58.4	2.7	67246	2	AC105035	AC105035 Homo sapi

ALIGNMENTS

RESULT 1
AR071760
LOCUS AR071760 2180 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 2 from patent US 5912142.
ACCESSION AR071760
VERSION AR071760.1 GI:7222648
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2180)
AUTHORS Kaufman,R.E. and Slentz-Kesler,K.
TITLE Gene product over expressed in cancer cells
JOURNAL Patent: US 5912142-A 2 15-JUN-1999;
FEATURES
Location/Qualifiers
source
1..2180
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match	100.0%	Score	2180;	DB	6;	Length	2180;
Best Local Similarity	100.0%	Pred. No.	0;	Mismatches	0;	Indels	0;
Matches	2180;	Conservative	0;				
Qy	1	ATTCTGCTTCTTTAGCGTGAACGCGGGTGGTTCCTCCGCGTGAATAATAATTCAC	60				
Db	1	ATTCTGCTTCTTTAGCGTGAACGCGGGTGGTTCCTCCGCGTGAATAATAATTCAC	60				
Qy	61	CGTACCGTTGTTGTGAACGCGGGTGGTTCCTCCGCGTGAATAATAATTCAC	120				
Db	61	CGTACCGTTGTTGTGAACGCGGGTGGTTCCTCCGCGTGAATAATAATTCAC	120				
Qy	121	CTCTTCTCATCTGCGAGGTGGTTCCTCCGCGTGAATAATAATTCAC	180				
Db	121	CTCTTCTCATCTGCGAGGTGGTTCCTCCGCGTGAATAATAATTCAC	180				
Qy	181	CCTCTCGTGTCTTTTCATTTCTCGGGGTCCGGGCGGAGAGTGCATCCAGAGGA	240				
Db	181	CCTCTCGTGTCTTTTCATTTCTCGGGGTCCGGGCGGAGAGTGCATCCAGAGGA	240				
Qy	241	CGGCGTCAGAGCGGACCCGGAGTGTTCAGAGCCAGTGCACAGGACGAGGCGCCA	300				
Db	241	CGGCGTCAGAGCGGACCCGGAGTGTTCAGAGCCAGTGCACAGGACGAGGCGCCA	300				

Db 241 GCGCGTCAGGAGCGGACCCGGAGTGTTCAGAGAGCCAGTGACAGAGACGAGGGGCCA 300
Qy 301 AGTCCACAGCCATGAGACCTGCCCTCCCTGGCATTCCTGGCCACAGCTTTCCAGGGCCCT 360
Db 301 AGTCCACAGCCATGAGACCTGCCCTCCCTGGCATTCCTGGCCACAGCTTTCCAGGGCCCT 360
Qy 361 TGGGACCCCTCCCTGTTTGGCTGCTCCTCTGAGTGCTCAGAAATGAAGCTGGGACAGCC 420
Db 361 TGGGACCCCTCCCTGTTTGGCTGCTCCTCTGAGTGCTCAGAAATGAAGCTGGGACAGCC 420
Qy 421 CATCTGCAAGAGGGGTAGTCTCTGTGTCTTGGGGCGAGACACCGCTCATGTCTGCAA 480
Db 421 CATCTGCAAGAGGGGTAGTCTCTGTGTCTTGGGGCGAGACACCGCTCATGTCTGCAA 480
Qy 481 CATCTCAACGCTTCTCCCATGTCAACATCAAGCTGGGTGCCACGGGCGAGGAGCGC 540
Db 481 CATCTCAACGCTTCTCCCATGTCAACATCAAGCTGGGTGCCACGGGCGAGGAGCGC 540
Qy 541 CATCTTCAATGAGTGGCTCCAGGTACTTCTCCGGGAGCGCTGGCAGCTCCAGTTCA 600
Db 541 CATCTTCAATGAGTGGCTCCAGGTACTTCTCCGGGAGCGCTGGCAGCTCCAGTTCA 600
Qy 601 GGGAGGGTGGACAGCTGGTGATCAAGCGCGCGGACTCCCATGCTGGGCTGTACAT 660
Db 601 GGGAGGGTGGACAGCTGGTGATCAAGCGCGCGGACTCCCATGCTGGGCTGTACAT 660
Qy 661 GTGGCACTCTGGGACACAGAGAAATAACAGACAAAGTCAAGCTGGAGGTTTCAAGTGC 720
Db 661 GTGGCACTCTGGGACACAGAGAAATAACAGACAAAGTCAAGCTGGAGGTTTCAAGTGC 720
Qy 721 AGAACCCAGTCCGCGCTGACACTGGGTCTGGCTGTGGCGGCTGGCAGCGGTGTCACTGCT 780
Db 721 AGAACCCAGTCCGCGCTGACACTGGGTCTGGCTGTGGCGGCTGGCAGCGGTGTCACTGCT 780
Qy 781 CTTTCATCTCTGTGCTCTGTGTCATGTTGGCTGTGGTACAGGTGCGCTGTGTCCAGCA 840
Db 781 CTTTCATCTCTGTGCTCTGTGTCATGTTGGCTGTGGTACAGGTGCGCTGTGTCCAGCA 840
Qy 841 ACGCGGGAGAGAAAGTTCTTCTCTAGAACCCAGATGAAGTGGCAGCCCTCAGAGC 900
Db 841 ACGCGGGAGAGAAAGTTCTTCTCTAGAACCCAGATGAAGTGGCAGCCCTCAGAGC 900
Qy 901 GGGAGCCAGCAGGCGCTGAGCAGAGCTCGCTGAACTGTGGACCCAGACTCCAGGC 960
Db 901 GGGAGCCAGCAGGCGCTGAGCAGAGCTCGCTGAACTGTGGACCCAGACTCCAGGC 960
Qy 961 CACCCAAAGCGCTGGCACTGGTGTCAAACCTCAACCTTGGAGCCCTGGAGCTGCT 1020
Db 961 CACCCAAAGCGCTGGCACTGGTGTCAAACCTTCAACCTTGGAGCCCTGGAGCTGCT 1020
Qy 1021 GTCCCCCAACCTTGTTCATATGCCGAGACCCATAGCCGCTGCAAGGCAGAG 1080
Db 1021 GTCCCCCAACCTTGTTCATATGCCGAGACCCATAGCCGCTGCAAGGCAGAG 1080
Qy 1081 GACACAGGAGCCAGCCCTGAGTCCGACCTTGGGTGGCGGCTGGGTCTCTCTGTC 1140
Db 1081 GACACAGGAGCCAGCCCTGAGTCCGACCTTGGGTGGCGGCTGGGTCTCTCTGTC 1140
Qy 1141 CACCGGGAGGACAGACCCGCTTGTGGAGGCTGGGCTCTGTGTCAACCACTCC 1200
Db 1141 CACCGGGAGGACAGACCCGCTTGTGGAGGCTGGGCTCTGTGTCAACCACTCC 1200
Qy 1201 TGGGTGCTGAGACCCCTTCCCTCCACCCCTCAGTCTTCCAGCTCTGTCTCAGT 1260
Db 1201 TGGGTGCTGAGACCCCTTCCCTCCACCCCTCAGTCTTCCAGCTCTGTCTCAGT 1260
Qy 1261 TTCCAAATGGAACCACTCACTCCGAGACCCGATTTACAGAGCATGCCCTCC 1320
Db 1261 TTCCAAATGGAACCACTCACTCCGAGACCCGATTTACAGAGCATGCCCTCC 1320
Qy 1321 CTCTGCCCTCATCAACCCAGACCCGAGCTCCCTTCTGTCACCCAGGCTGGTCCGG 1380
Db 1321 CTCTGCCCTCATCAACCCAGACCCGAGCTCCCTTCTGTCACCCAGGCTGGTCCGG 1380

Qy 1381 CCCCAGGTGTGGGTCCGCTCTCTCACTCCAGGCTCCCGGCCCAAGTGAAGGGGCC 1440
Db 1381 CCCCAGGTGTGGGTCCGCTCTCTCACTCCAGGCTCCCGGCCCAAGTGAAGGGGCC 1440
Qy 1441 CTGCCGAGCCCTCAGACACACTGGAGTTGAGGCTGGGGGGCCTTGGACATACCTGTC 1500
Db 1441 CTGCCGAGCCCTCAGACACACTGGAGTTGAGGCTGGGGGGCCTTGGACATACCTGTC 1500
Qy 1501 CTTTGGCTATGAGCAGGCTTGGGGCCCTCCGGGCGAGCCCGGGGGCCAGAGTAGGG 1560
Db 1501 CTTTGGCTATGAGCAGGCTTGGGGCCCTTCCGGGCGAGCCCGGGGGCCAGAGTAGGG 1560
Qy 1561 TCTGGGGCTTAGAGGCTGGGATGGCTCCTGGGCCCAACCGGCGGCGGCGAGGGCC 1620
Db 1561 TCTGGGGCTTAGAGGCTGGGATGGCTCCTGGGCCCAACCGGCGGCGGCGAGGGCC 1620
Qy 1621 GGGCTGGGAGGCGGCGGCGGCTGGGCTGGGGGTGAGTGGACGCTGCTCCGGGG 1680
Db 1621 GGGCTGGGAGGCGGCGGCGGCTGGGCTGGGGGTGAGTGGACGCTGCTCCGGGG 1680
Qy 1681 CTGGTCGCGCATCCCTCAGTCCCTCGGCCACCGGGGGTGGTCCCTGTCGCCACCGCA 1740
Db 1681 CTGGTCGCGCATCCCTCAGTCCCTCGGCCACCGGGGGTGGTCCCTGTCGCCACCGCA 1740
Qy 1741 CCTGCCAGGCTCTTTGGACCCAGATCTGTTTGTCTTGTCTTGTCTGTCAGTCCGGGG 1800
Db 1741 CCTGCCAGGCTCTTTGGACCCAGATCTGTTTGTCTTGTCTTGTCTGTCAGTCCGGGG 1800
Qy 1801 GCGCTTGTGCTTCTCATCTGTATGGGTGGAATAATCACCGGGAATCCCTTTCAGTTC 1860
Db 1801 GCGCTTGTGCTTCTCATCTGTATGGGTGGAATAATCACCGGGAATCCCTTTCAGTTC 1860
Qy 1861 TTTGAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAACAAACCGACTCACAACC 1920
Db 1861 TTTGAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAACAAACCGACTCACAACC 1920
Qy 1921 TCCAAAGTGTCCAAATGCAATTTTAAATGGAATAACAAATCTGAAAGAAACGCTTT 1980
Db 1921 TCCAAAGTGTCCAAATGCAATTTTAAATGGAATAACAAATCTGAAAGAAACGCTTT 1980
Qy 1981 TAGTGGCTTAAAGCCCAAAACGCTCCCTTAAGCGCTCTCGAGATGAAGACGGGGGGAGC 2040
Db 1981 TAGTGGCTTAAAGCCCAAAACGCTCCCTTAAGCGCTCTCGAGATGAAGACGGGGGGAGC 2040
Qy 2041 CCCAGCAGGTGAGAACCCCGCAGACCGGGGGGGCCCGGTGACCGAGGCTCGCACAG 2100
Db 2041 CCCAGCAGGTGAGAACCCCGCAGACCGGGGGGGCCCGGTGACCGAGGCTCGCACAG 2100
Qy 2101 CCGGCGCCCTGAGGCTGGGCGGAGCCAGGCTCCAGAGGGCGCGTGTGTCTCGGG 2160
Db 2101 CCGGCGCCCTGAGGCTGGGCGGAGCCAGGCTCCAGAGGGCGCGTGTGTCTCGGG 2160
Qy 2161 TTAAATAAGTTCGCTCCG 2180
Db 2161 TTAAATAAGTTCGCTCCG 2180

RESULT 2
AR097657
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

AR097657
Sequence 2 from patent US 6072034.
AR097657
AR097657.1 GI:12806387
Unknown.
Unclassified.
1 (bases 1 to 2180)
Kaufman, R.E. and Slentz-Kesler, K.
Gene product over expressed in cancer cells
Patent: US 6072034-A 2 06-JUN-2000;
Location/Qualifiers

linear
DNA
2180 bp
PAT 14-FEB-2001

[illegible]

Db	2041	CCCAGCAGGTGGAGACCCCGCAGGACCGCGCGCGCCCGGTGACCGAGGCGTCCGACAG	2100
Db	2041	CCCAGCAGGTGGAGACCCCGCAGGACCGCGCGCGCCCGGTGACCGAGGCGTCCGACAG	2100
Qy	2101	CGGCGCGCCCTGAGGCTGGGCGGCGGAGCCAGGCTCAAGAGGGCGCGTGTGTCTCGGG	2160
Db	2101	CGGCGCGCCCTGAGGCTGGGCGGCGGAGCCAGGCTCAAGAGGGCGCGTGTGTCTCGGG	2160
Qy	2161	TTAAATAAAGGTTCCGTCCG	2180
Db	2161	TTAAATAAAGGTTCCGTCCG	2180
RESULT 3	AR195192	2180 bp	DNA
LOCUS	Sequence 2 from patent US 6350615.		linear
DEFINITION	AR195192		
ACCESSION	AR195192.1	GI:20244629	
VERSION	Unknown.		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2180)		
AUTHORS	Kaufman, R.E. and Slentz-Keeler, K.		
TITLE	Gene product over expressed in cancer cells		
JOURNAL	Patent: US 6350615-A 2 26-FEB-2002;		
FEATURES	Location/Qualifiers		
source	1..2180		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN	Query Match 100.0%; Score 2180; DB 6; Length 2180;		
	Best Local Similarity 100.0%; Pred. No. 0;		
	Matches 2180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	ATTCTGCTTCCTTTAGCGTGAACCGGGTGGCGGTCCCGTGAATAATAAATTCAC	60
Db	1	ATTCTGCTTCCTTTAGCGTGAACCGGGTGGCGGTCCCGTGAATAATAAATTCAC	60
Qy	61	CGTCACGCTTGTGTGAACCGGGTGGTTCGGAACCTTGGAGGCTTCGTAACCCAG	120
Db	61	CGTCACGCTTGTGTGAACCGGGTGGTTCGGAACCTTGGAGGCTTCGGAACCCAG	120
Qy	121	CTCTCTCTCATCTGGAGGTGGTCCCGCGGGTCCCGCGCTTCCTCCCTGGCCCTC	180
Db	121	CTCTCTCTCATCTGGAGGTGGTCCCGCGGGTCCCGCGCTTCCTCCCTGGCCCTC	180
Qy	181	CCTCTCTGTCTTTCAATTTCTGGGGTCCCGGGCGGAGAGCTGCATCCAGAGGA	240
Db	181	CCTCTCTGTCTTTCAATTTCTGGGGTCCCGGGCGGAGAGCTGCATCCAGAGGA	240
Qy	241	CGCGTCCAGAGCGGACCGGAGGTGTTCAAGAGCCAGTCAAGAGCAGGAGGCGCA	300
Db	241	CGCGTCCAGAGCGGACCGGAGGTGTTCAAGAGCCAGTCAAGAGCAGGAGGCGCA	300
Qy	301	AGTCCACACCATGACAGACCTGCCCGCTGGCAATTCCTGGCCAGCTTCCAGGCGCT	360
Db	301	AGTCCACACCATGACAGACCTGCCCGCTGGCAATTCCTGGCCAGCTTCCAGGCGCT	360
Qy	361	TGGGACCCCTCTGTTTTGGTGGTCTCTGTTGTCTGGGGCGAGAACCCGTCATCTCTGCAA	480
Db	361	TGGGACCCCTCTGTTTTGGTGGTCTCTGTTGTCTGGGGCGAGAACCCGTCATCTCTGCAA	480
Qy	481	CATCTGCAAGAGGGGTAGTCTCTGTTGTCTGGGGCGAGAACCCGTCATCTCTGCAA	540
Db	481	CATCTGCAAGAGGGGTAGTCTCTGTTGTCTGGGGCGAGAACCCGTCATCTCTGCAA	540
Qy	541	CATCTCAATGAGGTGGCTCCAGGCTACTTCTCCCGGAGCGGTGGCGAGCTCCAGGTTCA	600
Db	541	CATCTCAATGAGGTGGCTCCAGGCTACTTCTCCCGGAGCGGTGGCGAGCTCCAGGTTCA	600

Db	1621	GGGCTGGGAGCGCGGGGGGGCTCGGGCTGGGGGGTCAAGTGGACGCTGCTCGGGG	1680
Qy	1681	CTGGTCGGCATCCCTCAGTCCCTCGGCCACCCCGGGGGTGGCTCCCTCGTGGCCACCGCA	1740
Db	1681	CTGGTCGGCATCCCTCAGTCCCTCGGCCACCCCGGGGGTGGCTCCCTCGTGGCCACCGCA	1740
Qy	1741	CTGCCGAGCCCTTTTGGACCCAGATCTGTTTCATGCTTTTGTCTTTCGTCTCTGCGGGG	1800
Db	1741	CTGCCGAGCCCTTTTGGACCCAGATCTGTTTCATGCTTTTGTCTTTCGTCTCTGCGGGG	1800
Qy	1801	GCCCTTTTGTCTTTCATCTGTATGGGGTGGAAAAATCACCGGGAATCCCCCTTCAGTTC	1860
Db	1801	GCCCTTTTGTCTTTCATCTGTATGGGGTGGAAAAATCACCGGGAATCCCCCTTCAGTTC	1860
Qy	1861	TTTGAAAAAGTTTCCATGACCTCGAATATCTGAAATGAAAGAAAAACAACCGACTCACAAACC	1920
Db	1861	TTTGAAAAAGTTTCCATGACCTCGAATATCTGAAATGAAAGAAAAACAACCGACTCACAAACC	1920
Qy	1921	TCCAAGTAGCTCCAATGCAATTTTAAAAATGAAAAACAAAAATCTGAAAGAAACGTCCT	1980
Db	1921	TCCAAGTAGCTCCAATGCAATTTTAAAAATGAAAAACAAAAATCTGAAAGAAACGTCCT	1980
Qy	1981	TAGTGGCTTTAAGCCCCCAACCGTCCCTTAAGGGCGTCTCGAGATGAAAGACGGGGGGAGC	2040
Db	1981	TAGTGGCTTTAAGCCCCCAACCGTCCCTTAAGGGCGTCTCGAGATGAAAGACGGGGGGAGC	2040
Qy	2041	CCAGCCAGGTGGAGACCCCGCAGACCGCGCGCGCGTGCACCGAGCCCTCGCACAG	2100
Db	2041	CCAGCCAGGTGGAGACCCCGCAGACCGCGCGCGCGTGCACCGAGCCCTCGCACAG	2100
Qy	2101	CCGCGCCGCTGAGGGTCGGGCGGAGCCAGGTCAGGGTCCAAAGAGGGGCGCTTTTGTCTCGGG	2160
Db	2101	CCGCGCCGCTGAGGGTCGGGCGGAGCCAGGTCAGGGTCCAAAGAGGGGCGCTTTTGTCTCGGG	2160
Qy	2161	TTAAAAAAGGTTCCGTCGG	2180
Db	2161	TTAAAAAAGGTTCCGTCGG	2180
RESULT 4			
CO776411			
LOCUS	CO776411	2000 bp	DNA linear PAT 11-MAR-2004
DEFINITION	Sequence 97 from Patent EP1394274.		
ACCESSION	CO776411		
VERSION	CO776411.1 GI:45379801		
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1		
AUTHORS	Ohtani, N., Sugita, Y., Yamaya, M., Kubo, H., Nagai, H. and Izuhara, K.		
TITLE	Methods of testing for bronchial asthma or chronic obstructive pulmonary disease		
JOURNAL	Patent: EP 1394274-A 97 03-MAR-2004;		
FEATURES	Genex Research, Inc. (JPN)		
	Location/Qualifiers		
source	1..2000		
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
ORIGIN			
Query Match 89.5%; Score 1951; DB 6; Length 2000;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;			
Qy	196	ATTTCCTGGGGCTCCGGGGCGCGGAAGCTGATCCAGAGAGGCGCGTCAGAGAGCG	255
Db	1	ATTTCCTGGGGCTCCGGGGCGCGGAAGCTGATCCAGAGAGGCGCGTCAGAGAGCG	60
Qy	256	GACCCGGAGTGTTCGAAGGCCAGTGCACAGGACCAAGGGGCCCAAGTCCCAACGACCAT	315

Wyeth (US); Burczynski, Michael E. (US); Twine, Natalie C. (US);
Dorner, Andrew J. (US); Trepicchio, William L. (US)
Location/Qualifiers
1. .2000
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES

source

ORIGIN

Query Match 89.5%; Score 1951; DB 6; Length 2000;

Best Local Similarity 99.8%; Pred. No. 0; Indels 3; Gaps 3;
Matches 1984; Conservative 0; Mismatches 0

QY	1396	CGCTCTCTCCACTCCAGGCTCCGCCCAAGTAGAGGGGCGCCCTCCGAGACCTCAG	1455
DB	1200	CGCTCTCTCCACTCCAGGCTCCGCCCAAGTAGAGGGGCGCCCTCCGAGACCTCAG	1259
QY	1456	ACACACTGAGATTACAGGCTGGGGGGCTTGGGCAATACCTGTCCCTTGGCTATGAGCA	1515
DB	1260	ACACACTGAGATTACAGGCTGGGGGGCTTGGGCAATACCTGTCCCTTGGCTATGAGCA	1319
QY	1516	GGCTTTGGGGGCTTCCGCGCAGCCCGGGGGCGGAGGTAGGTCTGGGGCTTAGAG	1575
DB	1320	GGCTTTGGGGGCTTCCGCGCAGCCCGGGGGCGGAGGTAGGTCTGGGGGCTTAGAG	1379
QY	1576	GCTGGGATGCTCTGGGCGCCACCGCCAGGGGGCAAGCGAGCGGGCTGGAGGCGGC	1635
DB	1380	GCTGGGATGCTCTGGGCGCCACCGCCAGGGGGCAAGCGAGCGGGCTGGAGGCGGC	1439
QY	1636	GGCGGGGCTCGGGCTGGGGGGTACAGTGAGCGCTGCCCTCCGGGCTGGTCGCGATCCC	1695
DB	1440	GGCGGGGCTCGGGCTGGGGGGTACAGTGAGCGCTGCCCTCCGGGCTGGTCGCGATCCC	1499
QY	1696	TGAGTCCCTCGGCCACCCCGGGGTCCCTCCCTCGTCCGCCACCGCACCTGCCGAGCTCTT	1755
DB	1500	TCAGTCCCTCGGCCACCCCGGGGTCCCTCCCTCGTCCGCCACCGCACCTGCCGAGCTCTT	1559
QY	1756	TGGACCCAGATCTGTTCATGCTTTTGTCTGTGTCACCTGGCGGGGGCCCTTTGATGCTT	1815
DB	1560	TGGACCCAGATCTGTTCATGCTTTTGTCTGTGTCACCTGGCGGGGGCCCTTTGATGCTT	1619
QY	1816	CATCTGTATGGGTGGGAAATACCGGGGATCCCGCTTCAGTTCTTTGAAAAAGTTCCA	1875
DB	1620	CATCTGTATGGGTGGGAAATACCGGGGATCCCGCTTCAGTTCTTTGAAAAAGTTCCA	1679
QY	1876	TGACTCGAATATCTGAATGAAGAAACAAACCGACTCAAACTCCAAAGTAGTCCAA	1935
DB	1680	TGACTCGAATATCTGAATGAAGAAACAAACCGACTCAAACTCCAAAGTAGTCCAA	1739
QY	1936	ATGCAATTTTAAATGGAAAAAAMAAATCTGAAGAAAGTCTTTAGTGGCTTTAAGCC	1995
DB	1740	ATGCAATTTTAAATGGAAAAAAMAAATCTGAAGAAAGTCTTTAGTGGCTTTAAGCC	1799
QY	1996	CCAAACGCTCCCTAAGCGCTCTCGAGATGAAGACGGGGGGGAG-CCCCAGCCAGGTGGA	2054
DB	1800	CCAAACGCTCCCTAAGCGCTCTCGAGATGAAGACGGGGGGGAGCCCGCCAGCGTGA	1859
QY	2055	GACCCCGCAGACCGCGCGCGCGCTGACCGAGGCTCGCACAGCCCGCGCCCTGAG	2114
DB	1860	GACCCCGCAGACCGCGCGCGCGCTGACCGAGGCTCGCACAGCCCGCGCCCTGAG	1919
QY	2115	GGTCGGGC-GAGCCAGGGTCCAGAGGGGCGCTTTGTCTCGGGTTAAATAAAGGTT	2173
DB	1920	GGTCGGGGCGGAGCGAGGCTCCAGAGGGGCGCTTTGTCTCGGGTTAAATAAAGGTT	1979
QY	2174	CCGTCCG 2180	
DB	1980	CGTCCG 1986	
RESULT 5			
CO861464			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
1			
Homo sapiens (human)			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
Burczynski, M., Twine, N., Dorner, A.J. and Trepicchio, W.L.			
METHODS FOR MONITORING DRUG ACTIVITIES IN VIVO /i			
Patent: WO 2004072265-A 97 26-AUG-2004			
JOURNAL			

FEATURES

source

ORIGIN

Query Match 89.5%; Score 1951; DB 6; Length 2000;

Best Local Similarity 99.8%; Pred. No. 0; Indels 3; Gaps 3;
Matches 1984; Conservative 0; Mismatches 0

QY	196	ATTTTCTCTGGGCTCCGGGGCGGAGAAAGCTGCATCCAGAGGAGCGGTCCAGAGCG	255
DB	1	ATTTTCTCTGGGCTCCGGGGCGGAGAAAGCTGCATCCAGAGGAGCGGTCCAGAGCG	60
QY	256	GACCCGGAGTGTTCAGAGCCAGTGAACAGACAGGGGCCAAGTCCACACGCCAT	315
DB	61	GACCCGGAGTGTTCAGAGCCAGTGAACAGACAGGGGCCAAGTCCACACGCCAT	120
QY	316	GCAGACCTGCCCTGGCATTCCTTGGCCACGTTTCCAGGGCCCTTGGGACCTCTCTGT	375
DB	121	GCAGACCTGCCCTGGCATTCCTTGGCCACGTTTCCAGGGCCCTTGGGACCTCTCTGT	180
QY	376	TTTGGCTGCCTCTTGAAGTGTCTCAGAAAGGCTGGGACAGCCCATCTGCACAGAGG	435
DB	181	TTTGGCTGCCTCTTGAAGTGTCTCAGAAAGGCTGGGACAGCCCATCTGCACAGAGG	240
QY	436	GTTAGTCTCTGTGTCTTGGGGCGAGAACACGTCATCTCTGCAACATCTTCCACGCTT	495
DB	241	GTTAGTCTCTGTGTCTTGGGGCGAGAACACGTCATCTCTGCAACATCTTCCACGCTT	300
QY	496	CTCCCATGTCAACATCAAGCTGCGTCCGACCGGACAGAGAGCCCATCTTCAATAGGT	555
DB	301	CTCCCATGTCAACATCAAGCTGCGTCCGACCGGACAGAGAGCCCATCTTCAATAGGT	360
QY	556	GGCTCCAGGCTACTTCTCCCGGACCGCTGGACCTCCAGTTCAGAGAGCGTGGACA	615
DB	361	GGCTCCAGGCTACTTCTCCCGGACCGCTGGACCTCCAGTTCAGAGAGCGTGGACA	420
QY	616	GCTGTGATCAAAAGCGCCCGGACTCCCATCTCTGGGCTGTACATGTGGCACTCTGGG	675
DB	421	GCTGTGATCAAAAGCGCCCGGACTCCCATCTCTGGGCTGTACATGTGGCACTCTGGG	480
QY	676	ACACAGAGAAATAACAGCAAGTCACTGGAGGTTTCAGGTGCAGAACCCAGTCCGC	735
DB	481	ACACAGAGAAATAACAGCAAGTCACTGGAGGTTTCAGGTGCAGAACCCAGTCCGC	540
QY	736	CCCTGACACTGGGTTCTGGCCCTGTCAGAGCGGTGTCTGCTGTCTTCTCTTGT	795
DB	541	CCCTGACACTGGGTTCTGGCCCTGTCAGAGCGGTGTCTGCTGTCTTCTCTTGT	600
QY	796	CGCTCTGGTCAATGTTCGCTGTGACAGGTGCGCTGTTCCTCCAGCAACCGCGGAGAA	855
DB	601	CGCTCTGGTCAATGTTCGCTGTGACAGGTGCGCTGTTCCTCCAGCAACCGCGGAGAA	660
QY	856	GTTCTTCTCTAGAACCCAGTGAAGTTCAGAGCGGTTCAGAGCGGAGCCAGCAGGG	915
DB	661	GTTCTTCTCTAGAACCCAGTGAAGTTCAGAGCGGTTCAGAGCGGAGCCAGCAGGG	720
QY	916	CCTGAGCAGAGCTCCGCTGAACTGTGGACCCCGACATCCGAGCCCAACCCAGGCGCT	975
DB	721	CCTGAGCAGAGCTCCGCTGAACTGTGGACCCCGACATCCGAGCCCAACCCAGGCGCT	780
QY	976	GGCACTGGTTCATAAACCCTTCCACTTGGAGCCCTTGGAGTGTCTGT-CCCCCAACCT	1035
DB	781	GGCACTGGTTCATAAACCCTTCCACTTGGAGCCCTTGGAGTGTCTGT-CCCCCAACCT	839
QY	1036	TGTTTCCATATGCGGACAGCCATAGCCGCTCTGAAGGAGGAGAGACACAGAGAGCCA	1095
DB	840	TGTTTCCATATGCGGACAGCCATAGCCGCTCTGAAGGAGGAGAGACACAGAGAGCCA	899

Qy	1096	GCCTTGATGCGGACCTTTGGGTGGCGGGGCTCGGTCTCTGTCGCACCCGGAGGGGACCA	1155
Db	900	GCCTTGATGCGGACCTTTGGGTGGCGGGGCTCGGTCTCTGTCGCACCCGGAGGGGACCA	959
Qy	1156	GACACGGCTTGCTTGGCAGGCTGGGCGCTCTGTGTGTCACCCACCTCTGGGTGCGTCAGAC	
Db	960	GACACGGCTTGCTTGGCAGGCTGGGCGCTCTGTGTGTCACCCACCTCTGGGTGCGTCAGAC	
Qy	1216	CCTTCCCTCCACCCCGCCAGGCTTCTCAAGCTCTGCTTCTCTCAGTTTTCAAAAATGGAACC	1275
Db	1020	CCTTCCCTCCACCCCGCCAGGCTTCTCAAGCTCTGCTTCTCTCAGTTTTCAAAAATGGAACC	1079
Qy	1276	ACCTCAGCTCGGACGACCCGACATTACAGGACGCAATGCCCTCTGCTGCTCATCAA	1335
Db	1080	ACCTCAGCTCGGACGACCCGACATTACAGGACGCAATGCCCTCTGCTGCTCATCAA	1139
Qy	1336	ACCACAGACCCGGAGCTCCCTTTCTGCCACCCACAGGCTGTGCGGGCCCAAGGTGTGGGT	1395
Db	1140	ACCACAGACCCGGAGCTCCCTTTCTGCCACCCACAGGCTGTGCGGGCCCAAGGTGTGGGT	1199
Qy	1396	CCGCTCTCTCCACTCCCAGGGGCTCCGCGCCCAAGTGAAGGGGGCCCTTGCOCGAGCCTCAG	1455
Db	1200	CCGCTCTCTCCACTCCCAGGGGCTCCGCGCCCAAGTGAAGGGGGCCCTTGCOCGAGCCTCAG	1259
Qy	1456	ACACATGAGATTACGGGCTGGGGGCGCTTGGGACATACCTGTCTTGCTTGCTGATGAGCA	1515
Db	1260	ACACATGAGATTACGGGCTGGGGGCGCTTGGGACATACCTGTCTTGCTTGCTGATGAGCA	1319
Qy	1516	GGCTTTGGGGGCGCTTCCGCGGACGACCCCGGGGGCGAGGTAGGGTCTGGGGGCTTAGAG	1575
Db	1320	GGCTTTGGGGGCGCTTCCGCGGACGACCCCGGGGGCGAGGTAGGGTCTGGGGGCTTAGAG	1379
Qy	1576	GCTGGGATGGCTCTGTGCCCCACCGCCAGGGGGCAAGCGCAGGCGCGGGCTTGGAGGCGGC	1635
Db	1380	GCTGGGATGGCTCTGTGCCCCACCGCCAGGGGGCAAGCGCAGGCGCGGGCTTGGAGGCGGC	1439
Qy	1636	GGCGGGGCTCGGGCTGGGGGTGAGTGAGACGTGCTCCGGGGCTGTGTGCGCATATCC	1695
Db	1440	GGCGGGGCTCGGGCTGGGGGTGAGTGAGACGTGCTCCGGGGCTGTGTGCGCATATCC	1499
Qy	1696	TCAGTCCCTCGGCCACCCCGGGGGTGGCTCCTCGTGCCACCGACCTGCGGAGCCTCTT	1755
Db	1500	TCAGTCCCTCGGCCACCCCGGGGGTGGCTCCTCGTGCCACCGACCTGCGGAGCCTCTT	1559
Qy	1756	TGGAACCCAGATCTGTTTCATGCTTTGTCTTCGTCACTGCGGGGGGCGCTTTTCATGTCTT	1815
Db	1560	TGGAACCCAGATCTGTTTCATGCTTTGTCTTCGTCACTGCGGGGGGCGCTTTTCATGTCTT	1619
Qy	1816	CATCTGTATGGGTGGAAAAATCACCGGGAATCCCGCTTCAGTCTTTTGAAAAAGTTCCA	1875
Db	1620	CATCTGTATGGGTGGAAAAATCACCGGGAATCCCGCTTCAGTCTTTTGAAAAAGTTCCA	1679
Qy	1876	TGACTCGAATATCTGAATGAAGAAACAAACCGACTCAAAACCTTCCAAAGTAGTCCAA	1935
Db	1680	TGACTCGAATATCTGAATGAAGAAACAAACCGACTCAAAACCTTCCAAAGTAGTCCAA	1739
Qy	1936	ATGCAATTTTTTAAATGCAAAAACAAAACTGTGAAAGAAACGTCTTTAGTGGCTTTAAGCC	1995
Db	1740	ATGCAATTTTTTAAATGCAAAAACAAAACTGTGAAAGAAACGTCTTTAGTGGCTTTAAGCC	1799
Qy	1996	CCAAAAAGCTCCCTAAGCGCTCTCGAGATGAAGACGCGGGGGAG-CCCCAGCCAGGTGGA	2054
Db	1800	CCAAAAAGCTCCCTAAGCGCTCTCGAGATGAAGACGCGGGGGAGCCCCAGCCAGGTGGA	1859
Qy	2055	GACCCCGCAGGACGCGGGGGCGCGGTGACCGAGGCTCGCACAGCCGCGCGCCCTGAG	2114
Db	1860	GACCCCGCAGGACGCGGGGGCGCGGTGACCGAGGCTCGCACAGCCGCGCGCCCTGAG	1919
Qy	2115	GGTCGGGCC--GAGCCAGGGTCCAAGAGGGGCGGTTTGTCTCTCGGGTTAAAAATAAGGTT	2173
Db	1920	GGTCGGGCCCGGAGCCAGGGTCCAAGAGGGGCGGTTTGTCTCTCGGGTTAAAAATAAGGTT	1979
Qy	2174	CCGTCOG 2180	

Db	1980	CCGTCGC 1986	
RESULT 6			
AX055560			
LOCUS	AX055560	2000 bp	DNA
DEFINITION	Sequence 3 from Patent WO073333.		
ACCESSION	AX055560		
VERSION	AX055560.1	GI:12228778	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Lyman, S.D. and Fauslow, W.C.		
TITLE	Ligand for cd7, and methods for use thereof		
JOURNAL	Patent: WO 007333-A 3 07-DEC-2000;		
	Immunex Corporation (US)		
FEATURES	Location/Qualifiers		
source	1..2000		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
CDS	119..865		
	/notes="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC22100.1"		
	/db_xref="GI:12228779"		
	/translation="MOTCLPAPGHVSQALGTLPLAASLAQNEGWDSPICTEGVVS		
	VSMGENTVMSCNISAFSHVNIKLRAHGQESALFNVEVAFGIFSRDGNQLQVQGGVAQL		
	VTKGARDHAGLYMHVLVGHQRNRRQVTLVEVGAQPSADPTGFVFPVAVTVAFVILL		
	VALVMPAWYRCRCSQQRREKFPFLPEPQKVAALRAQAQGLSRASAEMLTPDSEPTP		
	RPLALVFKPSPLGALLSPQLFPYAADP"		
ORIGIN			
Query Match	89.5%;	Score 1951;	DB 6; Length 2000;
Best Local Similarity	99.8%;	Pred. No. 0;	
Matches 1984; Conservative	0;	Mismatches	0; Indels 3; Gaps 3;
QY	196	ATTTTCTCTGGGCTCCGGGGCGGGAGAGAGTGCATCCACAGAGAGCGCGTCCAGAGCG	255
DB	1	ATTTTCTCTGGGCTCCGGGGCGGGAGAGTGCATCCACAGAGAGCGCGTCCAGAGCG	60
QY	256	GACCCGGGAGTGTTCAGAGCGAGTGCACAGAGCGAGGGCCCAAGTCCCACAGCCAT	315
DB	61	GACCCGGGAGTGTTCAGAGCGAGTGCACAGAGCGAGGGCCCAAGTCCCACAGCCAT	120
QY	316	GCAGACTGCCCTCGGCATTCCCTGGCCACGTTTCCAGGCCCTTGGGACCTCTCTGTT	375
DB	121	GCAGACTGCCCTCGGCATTCCCTGGCCACGTTTCCAGGCCCTTGGGACCTCTCTGTT	180
QY	376	TTTGGCTGCTCTCTCAGTGTCTCAGAAATGAAGCTGGGACAGGCCCATCTGCACAGAGG	435
DB	181	TTTGGCTGCTCTCTCAGTGTCTCAGAAATGAAGCTGGGACAGGCCCATCTGCACAGAGG	240
QY	436	GGTAGTCTCTGTCTTTGGGGCGAGAACACCGTCTGTCTGTCAAATCTTCCAAACGCTTT	495
DB	241	GGTAGTCTCTGTCTTTGGGGCGAGAACACCGTCTGTCTGTCAAATCTTCCAAACGCTTT	300
QY	496	CTCCCATGTCAAATCAAGTGTGCTGCCACCGGCGAGAGGCCCATCTTCAATGAGGT	555
DB	301	CTCCCATGTCAAATCAAGTGTGCTGCCACCGGCGAGAGGCCCATCTTCAATGAGGT	360
QY	556	GGCTCCAGGCTACTTCTCCGGGACGGCTGGCAGTCCAGGTTCCAGGAGGCGCTGGCACA	615
DB	361	GGCTCCAGGCTACTTCTCCGGGACGGCTGGCAGTCCAGGTTCCAGGAGGCGCTGGCACA	420
QY	616	GCTGGTGATCAAAAGGCGCCCGGAGTCCCATGTCTGGGTGTGATCATGTGGACACCTCTGGG	675
DB	421	GCTGGTGATCAAAAGGCGCCCGGAGTCCCATGTCTGGGTGTGATCATGTGGACACCTCTGGG	480

QY	676	ACACAGAGAAATAACAGACAAAGTCAACGCTGGAGGTTTCAGGTCCAGAACCCAGTCCGC	735	Db	1560	TGACCCAGATCTGTTCAATGCTTTTGTCTTCGTCACTGGCGGGGCCCTTTGATGCTT	1619
Db	481	ACACAGAGAAATAACAGACAAAGTCAACGCTGGAGGTTTCAGGTCCAGAACCCAGTCCGC	540	QY	1816	CATCTGTATGGGGTGGAATAATCACGGGAATCCCCCTTCAGTTCTTTTGAAAAGTTCGA	1875
QY	736	CCCTGACACTGGGTTCTGGCTGTGCCAGCGGTGTCACTGCTCTTCATCTCTTTGGT	795	Db	1620	CATCTGTATGGGGTGGAATAATCACGGGAATCCCCCTTCAGTTCTTTTGAAAAGTTCGA	1679
Db	541	CCCTGACACTGGGTTCTGGCTGTGCCAGCGGTGTCACTGCTCTTCATCTCTTTGGT	600	QY	1876	TGACTCGAATATCTGAAATGAAAGAAACAAACCGACTCAAAACCTCCAAAGTAGCTCAA	1935
QY	796	CGCTCTGGTTCATGTTTCGGCTGTGACAGTGGCGCTGTTCCAGCAACCCCGGAGAGAA	855	Db	1680	TGACTCGAATATCTGAAATGAAAGAAACAAACCGACTCAAAACCTCCAAAGTAGCTCAA	1739
Db	601	CGCTCTGGTTCATGTTTCGGCTGTGACAGTGGCGCTGTTCCAGCAACCCCGGAGAGAA	660	QY	1936	ATGCAATTTTTTAAATATGGAATAATCTGAAAGAAACGCTTTTAGTGGCTTTAAGCC	1995
QY	856	GTTCCTCTCTAGAACCCAGATGAAGTTCGACGCTTCAGAGCGGAGCCAGCAGGG	915	Db	1740	ATGCAATTTTTTAAATATGGAATAATCTGAAAGAAACGCTTTTAGTGGCTTTAAGCC	1799
Db	661	GTTCCTCTCTAGAACCCAGATGAAGTTCGACGCTTCAGAGCGGAGCCAGCAGGG	720	QY	1996	CCTAAACGTCCTTAAGCGCTTCGAGATGAAGACGGGGGGAG-CCCCAGGACAGGTGA	2054
QY	916	CTGAGCAGAGCTCCGCTGAAGTGTGACCCAGACTCCGAGCCACCCCAAGCCGCT	975	Db	1800	CCTAAACGTCCTTAAGCGCTTCGAGATGAAGACGGGGGGAGCCCCAGCAGGTGA	1859
Db	721	CTGAGCAGAGCTCCGCTGAAGTGTGACCCAGACTCCGAGCCACCCCAAGCCGCT	780	QY	2055	GACCCCGCAGGACGCGCGCGCGCTGACCGAGGCTTCGACAGCCGCGCCCTGAG	2114
QY	976	GGCACTGGTGTTCACAAACCTCACCACCTTGAGCCCTTGAGCTGTGTCCGCCCAACCT	1035	Db	1860	GACCCCGCAGGACGCGCGCGCGCTGACCGAGGCTTCGACAGCCGCGCCCTGAG	1919
Db	781	GGCACTGGTGTTCACAAACCTCACCACCTTGAGCCCTTGAGCTGTGTCCGCCCAACCT	839	QY	2115	GCTCGGCGCC-GAGCCAGGCTCCAGAGGGGCGGCTTTGTCTCGGCTTAAATAAGGTT	2173
QY	1036	TGTTTTCATATGCGCAGACCCATAGCCGCTGCAAGGACAGAGAGACACAGAGACCA	1095	Db	1920	GCTCGGCGCGGAGCCAGGGTCCAAAGAGGGCGGCTTTGTCTCGGCTTAAATAAGGTT	1979
Db	840	TGTTTTCATATGCGCAGACCCATAGCCGCTGCAAGGACAGAGAGACACAGAGACCA	899	QY	2174	CCGTCCG 2180	
QY	1096	GCCTGAGTCCGACCTTGGGTGGCGGGCTGGGTCTCTCTGCCACCCGAGGGCACA	1155	Db	1980	CCGTCCG 1986	
Db	900	GCCTGAGTCCGACCTTGGGTGGCGGGCTGGGTCTCTCTGCCACCCGAGGGCACA	959	RESULT 7			
QY	1156	GACACCGGCTTGCTGGAGGCTGGGCTCTGTGTCAACCACTCTGGGTGGTGACAG	1215	AX333755	2000 bp	DNA	linear
Db	960	GACACCGGCTTGCTGGAGGCTGGGCTCTGTGTCAACCACTCTGGGTGGTGACAG	1019	LOCUS	Sequence 4264 from Patent WO0194629.		
QY	1216	CTTTCCTCTCACCCTCCAGGCTTCCAGCTCTGCTCTCAGTTTCCAAATGAAC	1275	DEFINITION	AX333755		
Db	1020	CTTTCCTCTCACCCTCCAGGCTTCCAGCTCTGCTCTCAGTTTCCAAATGAAC	1079	ACCESSION	AX333755.1	GI:18124474	
QY	1276	ACCTCAGTCCGACGACCCGACTTACAGGAGCATGCCCCCTCTGCTCCCTCATCAA	1335	VERSION			
Db	1080	ACCTCAGTCCGACGACCCGACTTACAGGAGCATGCCCCCTCTGCTCCCTCATCAA	1139	KEYWORDS			
QY	1336	ACCACAGACCCGACTCCCTTCTGCAACCCAGGCTGCTCGGCCACAGGTGGGT	1395	SOURCE	Homo sapiens (human)		
Db	1140	ACCACAGACCCGACTCCCTTCTGCAACCCAGGCTGCTCGGCCACAGGTGGGT	1199	ORGANISM	Homo sapiens		
QY	1396	CGCTCTCTCAGTCCGAGGCTCCGCGCCCAAGTGAGGGGGCCCTTCGCGAGCCTCAG	1455	REFERENCE	Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,		
Db	1200	CGCTCTCTCAGTCCGAGGCTCCGCGCCCAAGTGAGGGGGCCCTTCGCGAGCCTCAG	1259	AUTHORS	Horrigan, S., Soppet, D.R. and Weaver, Z.		
QY	1456	ACACACTGAGTTACGGGCTGGGGGGCTTTGGACATACCTGCTCCCTTGCTATGAGCA	1515	TITLE	Cancer gene determination and therapeutic screening using signature		
Db	1260	ACACACTGAGTTACGGGCTGGGGGGCTTTGGACATACCTGCTCCCTTGCTATGAGCA	1319	JOURNAL	Patent: WO 0194629-A 4264 13-DEC-2001;		
QY	1516	GGCTTTGGGGGCTTCGCGGACGCCCCGAGGAGGAGTGGGTCTGGGGCTTAGAG	1575	FEATURES	Avalon Pharmaceuticals (US)		
Db	1320	GGCTTTGGGGGCTTCGCGGACGCCCCGAGGAGGAGTGGGTCTGGGGCTTAGAG	1379	source	1..2000		
QY	1576	GCTGGGATGGCTCTCGGCCCAACCGCCAGGGGGCAAGCGCAGCGGGCTGGAGGGCG	1635	source	/organism="Homo sapiens"		
Db	1380	GCTGGGATGGCTCTCGGCCCAACCGCCAGGGGGCAAGCGCAGCGGGCTGGAGGGCG	1439	source	/mol_type="unassigned DNA"		
QY	1636	GGCGCGGCTTCGGGCTGAGGTGAGCGCTGCTCGGGGCTGGTCCGATCC	1695	source	/db_xref="taxon:9606"		
Db	1440	GGCGCGGCTTCGGGCTGAGGTGAGCGCTGCTCGGGGCTGGTCCGATCC	1499	ORIGIN			
QY	1696	TCAGTCCCTCGGACACCGGGGGTCTGCTCCCTCGTGGCCACCGCACCTGCCAGCCTTT	1755	Query Match	89.5%;	Score 1951;	DB 6; Length 2000;
Db	1500	TCAGTCCCTCGGACACCGGGGGTCTGCTCCCTCGTGGCCACCGCACCTGCCAGCCTTT	1559	Best Local Similarity	99.8%;	Pred. No. 0;	
QY	1756	TGACCCAGATCTGTTTCATGCTTTTGTCTGCTACTGCGGGGGGCCCTTTGATGCTT	1815	Matches 1984;	Conservative	0; Mismatches	0; Indels 3; Gaps 3;

QY 376 TTGGCTGCTCTTGAAGTCTCAGAAATGAAGGCTGGGACAGCCCATCTGCACAGAGG 435
DB 181 TTGGCTGCTCTCTTGAAGTCTCAGAAATGAAGGCTGGGACAGCCCATCTGCACAGAGG 240
QY 436 GGTAGTCTCTGTGTCTTGGGGCGAGAAACCGTCTATGCTCTCAACATCTCCAAGCCCTT 495
DB 241 GGTAGTCTCTGTGTCTTGGGGCGAGAAACCGTCTATGCTCTCAACATCTCCAAGCCCTT 300
QY 496 CTCCATGTCAACATCAAGTCTGTCGCCCAACCGGCGAGGAGCGCCATCTTCAATGAGGT 555
DB 301 CTCCATGTCAACATCAAGTCTGTCGCCCAACCGGCGAGGAGCGCCATCTTCAATGAGGT 360
QY 556 GGTCTCAGGCTACTTCTCCGGGACGGCTGGCAGCTCCAGGTTCAAGGAGGCGTGGCACA 615
DB 361 GGTCTCAGGCTACTTCTCCGGGACGGCTGGCAGCTCCAGGTTCAAGGAGGCGTGGCACA 420
QY 616 GCTGTGTATCAAAAGGCGCCGGGACTCCCATGCTGGGCTGTATCATGTGTCACCTCGTGG 675
DB 421 GCTGTGTATCAAAAGGCGCCGGGACTCCCATGCTGGGCTGTATCATGTGTCACCTCGTGG 480
QY 676 ACACAGAGAAATAACAGAAAGTCAACGCTGGAGGTTTCAGGTGCAGAACCCAGTCCGC 735
DB 481 ACACAGAGAAATAACAGAAAGTCAACGCTGGAGGTTTCAGGTGCAGAACCCAGTCCGC 540
QY 736 CCTGACACTGGGTTCTGGCTGTGCGAGCGGTGTGCTACTGCTGTCTTCACTCTTGGT 795
DB 541 CCTGACACTGGGTTCTGGCTGTGCGAGCGGTGTGCTACTGCTGTCTTCACTCTTGGT 600
QY 796 CGCTCTGGTTCATGTTCCGCTGTACAGGTGCGCTGTTCCAGCAACCGCGGAGAGAA 855
DB 601 CGCTCTGGTTCATGTTCCGCTGTGAGGTGCGCTGTTCCAGCAACCGCGGAGAGAA 660
QY 856 GTTCTTCTCTTAGAACCCAGATGAAGTTCAGCGCTCAGCGCTCAGCGGGAGCCACAGGG 915
DB 661 GTTCTTCTCTTAGAACCCAGATGAAGTTCAGCGCTCAGCGGGAGCCACAGGG 720
QY 916 CTGAGCAGAGCCTCGCTGAACTGTGGAACCCAGACTCCGAGGCCACCCCAAGCGCGT 975
DB 721 CTGAGCAGAGCCTCGCTGAACTGTGGAACCCAGACTCCGAGGCCACCCCAAGCGCGT 780
QY 976 GGCACCTGGTTCAAACCTCACCCTTGGAGCCCTGGAGCTGTCTCCCCCAACCT 1035
DB 781 GGCACCTGGTTCAAACCTCACCCTTGGAGCCCTGGAGCTGTCTCCCCCAACCT 839
QY 1036 TGTTCCTATATCCGAGACCCATAGCCGCTGCAAGGCGAGAGGACACAGGAGGCA 1095
DB 840 TGTTCCTATATCCGAGACCCATAGCCGCTGCAAGGCGAGAGGACACAGGAGGCA 899
QY 1096 GGCCTGAGTCCGACCTTGGGTGGCGGGCTGGGTCTCTGTCGCCACCCGAGGGGACA 1155
DB 900 GGCCTGAGTCCGACCTTGGGTGGCGGGCTGGGTCTCTGTCGCCACCCGAGGGGACA 959
QY 1156 GACACGGCTGCTTGGCAGGCTGGGCTCTGTGTACCCACTCTGGGTGGTGCAGAC 1215
DB 960 GACACGGCTGCTTGGCAGGCTGGGCTCTGTGTACCCACTCTGGGTGGTGCAGAC 1019
QY 1216 CTTTCCCTTCCACCCCGCAGGCTCTTCCAGGCTCTGCTTCTCAGTTTCCAAAATGGAAC 1275
DB 1020 CTTTCCCTTCCACCCCGCAGGCTCTTCCAGGCTCTGCTTCTCAGTTTCCAAAATGGAAC 1079
QY 1276 ACCTCACTCTCCAGACCCGACTTACAGAGACGATGCCCCCTCTGTCCTCATCAA 1335
DB 1080 ACCTCACTCTCCAGACCCGACTTACAGAGACGATGCCCCCTCTGTCCTCATCAA 1139
QY 1336 ACCCAGACCCGGGACTCCCTTCTGCGACCCAGGCTGGTCCGCGCCAGGTGGGGT 1395
DB 1140 ACCCAGACCCGGGACTCCCTTCTGCGACCCAGGCTGGTCCGCGCCAGGTGGGGT 1199
QY 1396 CGGCTCTCTCCACTCCAGGCTCCGCGCCCAAGTGAGGGGCGCCCTGCGGAGGCTCAG 1455
DB 1200 CGGCTCTCTCCACTCCAGGCTCCGCGCCCAAGTGAGGGGCGCCCTGCGGAGGCTCAG 1259
QY 1456 ACACACTGGAGTTTCAAGGCTGGGGGGGCTTGGGCACATACCTGTCTCTTGGCTATGAGCA 1515

DB 1260 ACACACTGGAGTTTCAAGGCTGGGGGGGCTTGGGCACATACCTGTCTCTTGGCTATGAGCA 1319
QY 1516 GGTCTTGGGGGCTCTTCCGSGACGCCCGGGGGCCGAGGTAGGTCTGGGGGCTTAGAG 1575
DB 1320 GGTCTTGGGGGCTCTTCCGSGACGCCCGGGGGCCGAGGTAGGTCTGGGGGCTTAGAG 1379
QY 1576 GCTGGATGGCTCTTGGGCCCCCACCCAGGGGGCAAGCGACGCCGGGCTGGAGGCGGC 1635
DB 1380 GCTGGATGGCTCTTGGGCCCCCACCCAGGGGGCAAGCGACGCCGGGCTGGAGGCGGC 1439
QY 1636 GCGCGGCTCGGGCTGGGGGTGAGGTGAGCGCTCCCTCCGGGGCTGGTCCGGCATCCC 1695
DB 1440 GCGCGGCTCGGGCTGGGGGTGAGGTGAGCGCTCCCTCCGGGGCTGGTCCGGCATCCC 1499
QY 1696 TCAGTCCCTCGGCGACCCCGGGGTCGCTCCCTCGTGGCCACCGCACCTGCGGAGCCTTT 1755
DB 1500 TCAGTCCCTCGGCGACCCCGGGGTCGCTCCCTCGTGGCCACCGCACCTGCGGAGCCTTT 1559
QY 1756 TGGACCCAGATCTGTTCATGCTTTTGTCTTCTGTCTCACTGCGGGGGCCCTTTGATGTCTT 1815
DB 1560 TGGACCCAGATCTGTTCATGCTTTTGTCTTCTGTCTCACTGCGGGGGCCCTTTGATGTCTT 1619
QY 1816 CATCTGATGGGTGGAAAAATCACCGGGAATCCCTTTCAGTTCCTTTGAAAAAGTTCCA 1875
DB 1620 CATCTGATGGGTGGAAAAATCACCGGGAATCCCTTTCAGTTCCTTTGAAAAAGTTCCA 1679
QY 1876 TGACTCGAATATCTGAATGAAGAAAAAACAACCGACTCACAACTCTCAAGTAGTCCAA 1935
DB 1680 TGACTCGAATATCTGAATGAAGAAAAAACAACCGACTCACAACTCTCAAGTAGTCCAA 1739
QY 1936 ATGCAATTTTAAATGAAAAAACAATCTGAAAGAAACGCTTTTAGTGGCTTTAAGCC 1995
DB 1740 ATGCAATTTTAAATGAAAAAACAATCTGAAAGAAACGCTTTTAGTGGCTTTAAGCC 1799
QY 1996 CCAAAACCTCCTTAAGGCTCTCTGAGATGAAGACGGGGGGAG-CCCCAGCCAGGTGA 2054
DB 1800 CCAAAACCTCCTTAAGGCTCTCTGAGATGAAGACGGGGGGAGCCCCAGCCAGGTGA 1859
QY 2055 GACCCCGCAGACGCGCGCGCGCTGACGAGGCTCGCACAGCGCGGCCCTGAG 2114
DB 1860 GACCCCGCAGACGCGCGCGCGCTGACGAGGCTCGCACAGCGCGGCCCTGAG 1919
QY 2115 GGTGGGCC-GAGCCAGGTTCCAAAGAGGGGCGCTTTGTCTCGGTTAAAAAATAGGTT 2173
DB 1920 GGTGGGCCGAGGCCAGGTTCCAAAGAGGGGCGCTTTGTCTCGGTTAAAAAATAGGTT 1979
QY 2174 CGTCCG 2180
DB 1980 CCGTCCG 1986

RESULT 8
HSU77643
LOCUS HSU77643 2000 bp mRNA linear PRI 10-MAR-1998
DEFINITION Homo sapiens K12 protein precursor mRNA, complete cds.
ACCESSION U77643
VERSION U77643.1 GI:2062390
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
AUTHORS Slentz-Kesler, K.A., Hale, L.P. and Kaufman, R.E.
TITLE Identification and characterization of K12 (SCTM1), a novel human gene that encodes a Golgi-associated protein with transmembrane and secreted isoforms
JOURNAL Genomics 47 (3), 327-340 (1998)
MEDLINE 98149980
PUBMED 9480746
REFERENCE
AUTHORS Slentz-Kesler, K.A. and Kaufman, R.E.

TITLE	Direct Submission	Db
JOURNAL	Submitted (07-NOV-1996) Biochemistry, Duke University, Box 3250	Qy
FEATURES	DUMC, Durham, NC 27710, USA	Db
source	Location/Qualifiers	Qy
	1..2000	Db
	/organism="Homo sapiens"	Qy
	/mol_type="mRNA"	Db
	/db_xref="taxon:9606"	Qy
	/chromosome="17"	Db
	/map="17q25.2-25.3"	Qy
	/cell_type="K562 erythroleukemic"	Db
CDS	119..865	Qy
	/note="type Ia transmembrane protein"	Db
	/codon_start=1	Qy
	/product="K12 protein precursor"	Db
	/protein_id="AAC52044.1"	Qy
	/db_xref="GI:2062391"	Db
	/translation="MQTCFLAPPHVSHQALGTLFLAASLSAQNEGWDSPICTEGVSVS VSGENTVMSCNISAFSHVINIKLRAHGEAIFNEVAPGFSDRGWLQVQGGVAOL VIKGARDSHAGIYMHVLVGHQRNRRQVLEVSAGPQAPDPTGFMPVPAVITAFILL VALVMFAWYRCRCSQORREKFFLLEPQMKVAALRAGAQQGLSRASBELWTPDSEPTP RPLALVPKPSPLGALELLSPQLFPVADP"	Qy
sig_peptide	119..202	Db
mat_peptide	203..862	Qy
misc_feature	551..623	Db
misc_feature	752..862	Qy
	/note="encodes transmembrane domain"	Db
	/note="encodes proline-rich region"	Qy
ORIGIN		Db
Query Match	89.5%; Score 1951; DB 9; Length 2000;	
Best Local Similarity	99.8%; Pred. No. 0;	
Matches 1984; Conservative	0; Mismatches 0; Indels 3; Gaps 3;	
Qy	196 ATTTTCCTGGGCTCGGGCGGGAGAGCTGCATCCAGAGAGCGGCTCCAGAGCG 255	
Db	1 ATTTTCCTGGGCTCGGGCGGGAGAGCTGCATCCAGAGAGCGGCTCCAGAGCG 60	
Qy	256 GACCCGGGAGTGTTCAGAGCCAGTGCACAGAGCCAGGGGCCCAAGTCCACACCAT 315	
Db	61 GACCCGGGAGTGTTCAGAGCCAGTGCACAGAGCCAGGGGCCCAAGTCCACACCAT 120	
Qy	316 GCAGACCTGCCCTTGGCATTCCTGGCCAGCTTCCAGGCGCTTGGGACCTCCTGTT 375	
Db	121 GCAGACCTGCCCTTGGCATTCCTGGCCAGCTTCCAGGCGCTTGGGACCTCCTGTT 180	
Qy	376 TTGGCTGCTCCTTGAAGTGTCTAGAGTGAAGGCTGGGACAGGCCCATCTGCACAGAGG 435	
Db	181 TTGGCTGCTCCTTGAAGTGTCTAGAGTGAAGGCTGGGACAGGCCCATCTGCACAGAGG 240	
Qy	436 GGTAGTCTCTGTGTTTGGGCGGAGAACACCGTGTATCTCTGCACATCTCCACGCTT 495	
Db	241 GGTAGTCTCTGTGTTTGGGCGGAGAACACCGTGTATCTCTGCACATCTCCACGCTT 300	
Qy	496 CTCACATCTCAACATCAAGTCCGTCGCGGAGAGAGCGCCCATCTTCAATGAGGT 555	
Db	301 CTCACATCTCAACATCAAGTCCGTCGCGGAGAGAGCGCCCATCTTCAATGAGGT 360	
Qy	556 GGCTCCAGGCTACTTCTCCCGGAGCGGTGGCAGCTCCAGTTTCAGGGAGCGTGGCACA 615	
Db	361 GGCTCCAGGCTACTTCTCCCGGAGCGGTGGCAGCTCCAGTTTCAGGGAGCGTGGCACA 420	
Qy	616 GTGTGTATCAAGAGCGCGGAGCTCCCATGCTGGGCTGTACATGTGGACCTCGTGGG 675	
Db	421 GTGTGTATCAAGAGCGCGGAGCTCCCATGCTGGGCTGTACATGTGGACCTCGTGGG 480	
Qy	676 ACACAGAGAAATAACAGACAAAGTCAAGTGGAGGTTTCAGTGCAGAAACCCAGTCCGC 735	
Db	481 ACACAGAGAAATAACAGACAAAGTCAAGTGGAGGTTTCAGTGCAGAAACCCAGTCCGC 540	
Qy	736 CCTGTGACTGGGTTCTGGCCGTGGCCAGCGGTGTCTGCTGCTTCATCCTCTTGGT 795	

541	CCCTGACACTGGGTTCTGGCCTGTGCCAGCGGTGTCTACTGTCTTCTTCTTCTTGGT 600	Db
796	CGCTCTGGTTCATGTTTCGCTGTACAGGTGCGCTGTTCCTCCAGCAACGCGCGGAGAGAA 855	Qy
601	CGCTCTGGTTCATGTTTCGCTGTACAGGTGCGCTGTTCCTCCAGCAACGCGCGGAGAGAA 660	Db
856	GTTCTTCTCTCTAGAAACCCAGATGAAGTCCAGCCCTCAGAGCGGGAGCCAGCAGGG 915	Qy
661	GTTCTTCTCTCTAGAAACCCAGATGAAGTCCAGCCCTCAGAGCGGGAGCCAGCAGGG 720	Db
916	CTTGAGAGAGCCTCGCTGAACTGTGGACCCAGACTCCGAGCCACCCCAAGGCGCT 975	Qy
721	CTTGAGAGAGCCTCGCTGAACTGTGGACCCAGACTCCGAGCCACCCCAAGGCGCT 780	Db
976	GGCACTGGTGTCAAAACCCCTCACCACTTGGAGCCCTGGAGCTGTCTCTCCCCCAACCT 1035	Qy
781	GGCACTGGTGTCAAAACCCCTCACCACTTGGAGCCCTGGAGCTGTCTCTCTCCCAACCT 839	Db
1036	TGTTTTCATATGCCGAGACCCCATAGCCGCTCTCAAGGCGAGAGACACAGAGAGCCA 1095	Qy
840	TGTTTTCATATGCCGAGACCCCATAGCCGCTCTCAAGGCGAGAGACACAGAGAGCCA 899	Db
1096	GCCTGTAGTCCGACCTTGGGTGGCGGGCTGGGTCTCTCTGTCCTCCACCCGAGAGGACA 1155	Qy
900	GCCTGTAGTCCGACCTTGGGTGGCGGGCTGGGTCTCTCTGTCCTCCACCCGAGAGGACA 959	Db
1156	GACACCGGCTTGTGGCAGGCTGGGCTCTGTGTACCACTCTCTGGGTGCTGTGAGAC 1215	Qy
960	GACACCGGCTTGTGGCAGGCTGGGCTCTGTGTACCACTCTCTGGGTGCTGTGAGAC 1019	Db
1216	CCTTCCCTCTCACCCCGAGCTTTCAGAGTCTGCTTCTCAGTTTCCAAATATGGAAC 1275	Qy
1020	CCTTCCCTCTCACCCCGAGCTTTCAGAGTCTGCTTCTCAGTTTCCAAATATGGAAC 1079	Db
1276	ACCTCACCTCCGAGACCCCGACTTACAGAGCCATGCCCTCTCTCTGCTGCTCATCAA 1335	Qy
1080	ACCTCACCTCCGAGACCCCGACTTACAGAGCCATGCCCTCTCTCTGCTGCTCATCAA 1139	Db
1336	ACCCACAGACCCGAGCTCTCTTTCGCCAAGCTGCTGCTGCGGCGGCTGCTGGGT 1395	Qy
1140	ACCCACAGACCCGAGCTCTCTTTCGCCAAGCTGCTGCTGCGGCGGCTGCTGGGT 1199	Db
1396	CGGCTCTCTCACCTCCAGAGGCTCGCGGCCCAAGTGAAGGGGCGCTGCGGAGCCTCAG 1455	Qy
1200	CGGCTCTCTCACCTCCAGAGGCTCGCGGCCCAAGTGAAGGGGCGCTGCGGAGCCTCAG 1259	Db
1456	ACACACTGGAGTTCAAGGCTGGGGGGCTTGGGACATACCTGTCTCTTGGCTATGAGCA 1515	Qy
1260	ACACACTGGAGTTCAAGGCTGGGGGGCTTGGGACATACCTGTCTCTTGGCTATGAGCA 1319	Db
1516	GGCTTTGGGGGCTTTCGCGGCGAGCCCGGGGGCGGAGGTAGGTCTGGGGGCTTAGAG 1575	Qy
1320	GGCTTTGGGGGCTTTCGCGGCGAGCCCGGGGGCGGAGGTAGGTCTGGGGGCTTAGAG 1379	Db
1576	GCTGGATGGCTCTTGGGCGCCACCGCCAGGGGGCAAGCGCAGGCGGGCTGGAGGCGGC 1635	Qy
1380	GCTGGATGGCTCTTGGGCGCCACCGCCAGGGGGCAAGCGCAGGCGGGCTGGAGGCGGC 1439	Db
1636	GGCGGGGCTCGGGGTGGGGGTCAAGTGGAGCGCTGCTCCGGGGCTGTCGCGCATCC 1695	Qy
1440	GGCGGGGCTCGGGGTGGGGGTCAAGTGGAGCGCTGCTCCGGGGCTGTCGCGCATCC 1499	Db
1696	TGAGTCCCTGGGCGACCCCGGGGGTGGCTCTCTGTCGCCACCGCACCTGCGGAGCTCTT 1755	Qy
1500	TGAGTCCCTGGGCGACCCCGGGGGTGGCTCTCTGTCGCCACCGCACCTGCGGAGCTCTT 1559	Db
1756	TGAGTCCCTGGGCGACCCCGGGGGTGGCTCTCTGTCGCCACCGCACCTGCGGAGCTCTT 1815	Qy
1560	TGAGTCCCTGGGCGACCCCGGGGGTGGCTCTCTGTCGCCACCGCACCTGCGGAGCTCTT 1619	Db
1816	CATCTGTATGGGGTGGAAAAATCACCGGGAATCCCTCTTCAAGTTCTTTGAAAAAGTTCCA 1875	Qy
1620	CATCTGTATGGGGTGGAAAAATCACCGGGAATCCCTCTTCAAGTTCTTTGAAAAAGTTCCA 1679	Db

361 GCTGGCAGCTCCAGTTTCAGGAGGCGGTGGCCACAGCTGGTGTATCAAAAGGCGCCCGGACT 420
Db
642 CCCATGCTGGGCTGTACATGTGGACACTCTGTTGGGACACACAGAGAAATAACAGACAACTCA 701
Qy
421 CCCATGCTGGGCTGTACATGTGGACACTCTGTTGGGACACACAGAGAAATAACAGACAACTCA 480
Db
702 CGCTGGAGTTTCAGTGCAGAAACCCAGTCCGCCCTTGACACTGGGTTCTGGGCTGTGC 761
Qy
481 CGCTGGAGTTTCAGTGCAGAAACCCAGTCCGCCCTTGACACTGGGTTCTGGGCTGTGC 540
Db
762 CAGCGTGTCTACTGCTGTTTCATCTCTTGGTGGCTCTGGTCAATGTTCCCTGGGTACA 821
Qy
541 CAGCGTGTCTACTGCTGTTTCATCTCTTGGTGGCTCTGGTCAATGTTCCCTGGGTACA 600
Db
822 GGTGGCGTGTCTCCAGCAACCCGAGGAGAAAGTCTCTCTCTAGAACCCCGAGTGA 881
Qy
601 GGTGGCGTGTCTCCAGCAACCCGAGGAGAAAGTCTCTCTCTAGAACCCCGAGTGA 660
Db
882 AGTTCGAGCCCTTCAGAGCGGAGCCACAGAGGCTGTAGAGAGCTCCGCTGAACGTGT 941
Qy
661 AGTTCGAGCCCTTCAGAGCGGAGCCACAGAGGCTGTAGAGAGCTCCGCTGAACGTGT 720
Db
942 GGAACCCAGACTCCGAGCCACCCCAAGCGCTGGCACTGGTGTCAAAACCTCCACCAC 1001
Qy
721 GGAACCCAGACTCCGAGCCACCCCAAGCGCTGGCACTGGTGTCAAAACCTCCACCAC 780
Db
1002 TTGGAGCCCTGGAGTGTCTCCGCCCAACCCCTGTTTCCATATAGTCGCGAGACCATAG 1061
Qy
781 TTGGAGCCCTGGAGTGTCTCCGCCCAACCCCTGTTTCCATATAGTCGCGAGACCATAG 839
Db
1062 CCGCTTCAGAGGAGAGACACAGAGAGCCGCTGAGTGGCGACTTGGGTGGG 1121
Qy
840 CCGCTTCAGAGGAGAGACACAGAGAGCCGCTGAGTGGCGACTTGGGTGGG 899
Db
1122 GGGCTGGTCTCTGCTCCACCGAGGAGCAGACACCGGCTTCTTGGCAGGCTGG 1181
Qy
900 GGGCTGGTCTCTGCTCCACCGAGGAGCAGACACCGGCTTCTTGGCAGGCTGG 959
Db
1182 CCTGTGTATACCACTCTGCTGGGTGGTGGAGACCTTCCCTCCACCCCGAGTCTTC 1241
Qy
960 CCTGTGTATACCACTCTGCTGGGTGGTGGAGACCTTCCCTCCACCCCGAGTCTTC 1019
Db
1242 CAAGCTCTGCTCTCTGCTTCCAAATGGAACCACTTCACTCCGCGAGACCCGACTTA 1301
Qy
1020 CAAGCTCTGCTCTCTGCTTCCAAATGGAACCACTTCACTCCGCGAGACCCGACTTA 1079
Db
1302 CCAGGACGATGCGCCCTCTGCTGCTTCAATCAAAACCAAGAGCCGACTTCTCTG 1361
Qy
1080 CCAGGACGATGCGCCCTCTGCTGCTTCAATCAAAACCAAGAGCCGACTTCTCTG 1139
Db
1362 CCAGGACGATGCGCCCTCTGCTGCTTCAATCAAAACCAAGAGCCGACTTCTCTG 1421
Qy
1140 CCAGGACGATGCGCCCTCTGCTGCTTCAATCAAAACCAAGAGCCGACTTCTCTG 1199
Db
1422 CCAGGACGATGCGCCCTCTGCTGCTTCAATCAAAACCAAGAGCCGACTTCTCTG 1479
Qy
1200 CCAGGACGATGCGCCCTCTGCTGCTTCAATCAAAACCAAGAGCCGACTTCTCTG 1259
Db
1480 GGGCTTGGCACAATACCTGCTGCTGCTTCAATCAAAACCAAGAGCCGACTTCTCTG 1539
Qy
1260 GGGCTTGGCACAATACCTGCTGCTGCTTCAATCAAAACCAAGAGCCGACTTCTCTG 1319
Db
1540 GGGCTTGGCACAATACCTGCTGCTGCTTCAATCAAAACCAAGAGCCGACTTCTCTG 1599
Qy
1320 GGGCTTGGCACAATACCTGCTGCTGCTTCAATCAAAACCAAGAGCCGACTTCTCTG 1378
Db
1600 GGGCTTGGCACAATACCTGCTGCTGCTTCAATCAAAACCAAGAGCCGACTTCTCTG 1659
Qy
1379 GGGCTTGGCACAATACCTGCTGCTGCTTCAATCAAAACCAAGAGCCGACTTCTCTG 1437
Db
1660 AGGTGGAGCGCTG-CCTTCGGGGGCTGGTTCGGGAGCTTCCCTCAGTCCCTCGGCGCACCCGGGG 1718
Qy
1438 AGGTGGAGCGCTG-CCTTCGGGGGCTGGTTCGGGAGCTTCCCTCAGTCCCTCGGCGCACCCGGGG 1497
Db

1719 TGGCTCCCTGTCGCCACCGACCTGCGAGCCTCTTTGGACCCAGATCTGTTATGCTT 1778
Qy
1498 TGGCTCCCTGTCGCCACCGACCTGCGAGCCTCTTTGGACCCAGATCTGTTATGCTT 1557
Db
1779 TTGTCTTCCTCACTGCGGGGGCCCTTTGATGTTCTTCACTGTATGGGTGGAAAAATC 1838
Qy
1558 TTGTCTTCCTCACTGCGGGGGCCCTTTGATGTTCTTCACTGTATGGGTGGAAAAATC 1617
Db
1839 ACCGGAAATCCCTTCAGTTCTTTGAAAAGTCCATGACTCGAATATCTGAAATGAAG 1898
Qy
1618 ACCGGAAATCCCTTCAGTTCTTTGAAAAGTCCATGACTCGAATATCTGAAATGAAG 1677
Db
1899 AAAACAAACCGACTCACAACCTCCAAAGTAGCTCCAAATGCAATTTTAAAATGAAAAC 1958
Qy
1678 AAAACAAACCGACTCACAACCTCCAAAGTAGCTCCAAATGCAATTTTAAAATGAAAAC 1737
Db
1959 AAAAATCTGAAAGAA 1974
Qy
1738 AAAAATCTGAAAGAA 1753
Db

RESULT 10

AC132938 154252 bp DNA linear PRI 20-FBB-2003
LOCUS Homo sapiens chromosome 17, clone RP13-20L14, complete sequence.
DEFINITION AC132938
ACCESSION AC132938.9 GI:28436295
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 154252)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 17, clone RP13-20L14
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 154252)
AUTHORS Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
Camurata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Melrim, J., Meneus, L., Minova, T., Miura, C., Norman, C.H.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Peterson, K.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Rise, C., Rogov, P.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Seaman, S., Severy, P.,
Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission
Submitted (04-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 154252)
AUTHORS Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhgalter, B., Camurata, J., Chang, J., Choepel, Y.,
Collymore, A., Cooke, A., Cooke, P., Corum, B., DeArrellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-pierre, N., Hafez, N., Hagoopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Melrim, J., Meneus, L., Minova, T., Miura, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,

O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vasiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (31-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 154252)

Biren, B., Nuebaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouktgatier, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Katat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vasiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (20-FEB-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 20, 2003 this sequence version replaced gl:28173196.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)

Only the first 154.3 kilobases of this clone are being submitted. The remainder overlaps accession number AC124287 [WICGR project L27363].

FEATURES	source
02756331.	Location/Qualifiers
1..154252	
/organism="Homo sapiens"	
/mol_type="genomic DNA"	
/db_xref="taxon:9606"	
/chromosome="17"	
/map="17"	
/clone="RP13-20L14"	
/clone_lib="RPC1-13 Human Female BAC"	
1..32	
/rpt_family="Alu"	
167..473	
/rpt_family="AluSx"	
490..727	
/rpt_family="L1ME"	
1062..1336	
/rpt_family="AluUb"	
161..1974	
/rpt_family="MSTA"	
3544..3664	
/rpt_family="L12"	
repeated region	complement(3902..3954)

repeat_region	/rpt_family="MER5A" complement(4782..4945)
repeat_region	/rpt_family="MIR" 5143..5294
repeat_region	/rpt_family="MIR" complement(5298..5589)
repeat_region	/rpt_family="AlusX" 5696..6016
repeat_region	/rpt_family="AluJb" complement(6156..6436)
repeat_region	/rpt_family="AluY" complement(6445..6733)
repeat_region	/rpt_family="AlusX" complement(7025..7336)
repeat_region	/rpt_family="AlusX" 9135..9432
repeat_region	/rpt_family="MER72" 9440..9721
repeat_region	/rpt_family="MER72" complement(9727..9981)
repeat_region	/rpt_family="LTR8A" complement(9882..10192)
repeat_region	/rpt_family="AluSg" complement(10193..10749)
repeat_region	/rpt_family="LTR8A" complement(10750..11048)
repeat_region	/rpt_family="AluJb" complement(11049..11084)
repeat_region	/rpt_family="LTR8A" 11108..11399
repeat_region	/rpt_family="AlusX" 11598..11912
repeat_region	/rpt_family="AluSg" 11946..12247
repeat_region	/rpt_family="AlusX" complement(12259..12412)
repeat_region	/rpt_family="L1MC5" 12462..13718
repeat_region	/rpt_family="SVA" 13781..14081
repeat_region	/rpt_family="AlusX" complement(14198..14852)
repeat_region	/rpt_family="MER4D" complement(14853..15162)
repeat_region	/rpt_family="AluY" complement(15163..15344)
repeat_region	/rpt_family="MER4D" complement(15345..15845)
repeat_region	/rpt_family="L1MC5" 16100..16120
repeat_region	/rpt_family="GC_rich" 16155..16177
repeat_region	/rpt_family="GC_rich" complement(16624..16919)
repeat_region	/rpt_family="AluJb" complement(18971..19105)
repeat_region	/rpt_family="FLAM_A" complement(19548..19847)
repeat_region	/rpt_family="AluYa5" 23709..23833
repeat_region	/rpt_family="FLAM_C" 24811..24899
repeat_region	/rpt_family="MIR" 25425..25642
repeat_region	/rpt_family="C-rich" 27854..27907
repeat_region	/rpt_family="MIR" 28386..28669
repeat_region	/rpt_family="MER41D" complement(28670..28964)
repeat_region	/rpt_family="AlusX" 28965..29222
repeat_region	/rpt_family="MER41D"

[illegible]

Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Nguyen,C., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Roman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Norman,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (10-DEC-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 171569)

Birren,B., Nuebaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (14-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 171569)

Birren,B., Nuebaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Meneus,L., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (02-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
6 (bases 1 to 171569)

Birren,B., Nuebaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,

Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (20-MAY-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 20, 2003 this sequence version replaced gi:29469601.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L28146
Center clone name: 516_M_14

FEATURES
source
1..171569
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP13-516M14"
/clone_lib="RPC1-13 Human Female BAC"
complement(1..74)
/rpt_family="Alu"
complement(520..1125)
/rpt_family="PABL_A"
1768..2096
/rpt_family="AluX"
complement(2186..2397)
/rpt_family="MER66A"
complement(2403..3174)
/rpt_family="MER83A-int"
complement(3768..4157)
/rpt_family="LTR38B"
complement(4159..4439)
/rpt_family="AluX"
4459..4536
/rpt_family="HUERS-P3"
4537..4831
/rpt_family="AluY"
4832..5078
/rpt_family="HUERS-P3"
5079..5377
/rpt_family="AluY"
5378..5490
/rpt_family="HUERS-P3"
5491..5797
/rpt_family="G-rich"
5815..5999
/rpt_family="G-rich"
5925..6094
/rpt_family="CAGGG)n"
complement(6626..7181)
/rpt_family="MER41E"
complement(7371..7679)
/rpt_family="AluJb"
7716..8020
/rpt_family="AluSg"
8322..8480
/rpt_family="FAM"

QY	1666	ACGCTG-CCTCGGGGCTGGTGGCGCATCCCTCAGTCCCTCGGCACCCCGGGGTGCCTC	1724
DB	24521	ACGCGCCCTCCGGGGCTGGAGCGGCATCCCTCAGTCCCTCGGCACCCCGGGGTGCCTC	24580
QY	1725	CCTCGTGGCCACCGCACCTGCGGAGCTCTTTGACACCAGATCTGTTCATGCTTTTGTCT	1784
DB	24581	CCTCGTGGCCACCGCACCTGCGGAGCTCTTTGAGCCCTTTTGGACCCAGATCTGTTCATGCTTTTGTCT	24640
QY	1785	TGCTCACTCGCGCGGGGCCCTTTGATGCTTTCATCTGTATCGGGTGGAAAAATCACCGG	1844
DB	24641	TGCTCACTCGCGCGGGGCCCTTTGATGCTTTCATCTGTATCGGGTGGAAAAATCACCGG	24700
QY	1845	AATCCCCCTCAGTCTTTTGAAGAAAGTTCATGACTCGAAATATCTGAAATGAAGAAACA	1904
DB	24701	AATCCCCCTCAGTCTTTTGAAGAAAGTTCATGACTCGAAATATCTGAAATGAAGAAACA	24760
QY	1905	AACCGACTCACAAACCTCCAGTAGCTCCAAATGCAATTTTAAATGGAAAAACAAAT	1964
DB	24761	AACCGACTCACAAACCTCCAGTAGCTCCAAATGCAATTTTAAATGGAAAAACAAAT	24820
QY	1965	CTGAAGAAACGCTTTTAGTGCGCTTTAAGCCCCAAAAGCTCCCTAAGCGCTCTCGAGAT	2024
DB	24821	CTGAAGAAACGCTTTTAGTGCGCTTTAAGCCCCAAAAGCTCCCTAAGCGCTCTCGAGAT	24880
QY	2025	GAAGACGGGGGGGAG-CCCCAGCCAGGTGGAGACCCCGCA-GGACGCGCGCGCGCCCGGT	2082
DB	24881	GAAGACGGGGGGGAGCCCCCAGCAGGTGGAGACCCCGCAGGAGCGCGCGCGCCCGGT	24940
QY	2083	GACGAGGCTCCGACAGCGCGCGCCCTCAGGCTCGGGCC-GAGCAGGCGTCCAAGAGG	2141
DB	24941	GACGAGGCTCCGACAGCGCGCGCCCTCAGGCTCGGGCCGCGCGGAGCCAGGCTCCAAGAGG	25000
QY	2142	GGCGGTTTGTGCTCGGGTAAATAAAGTTCGCTCCG	2180
DB	25001	GGCGGTTTGTGCTCGGGTAAATAAAGTTCGCTCCG	25039
RESULT 12			
LOCUS	BD270355	1464 bp	linear
DEFINITION	50 human secreted proteins.		
ACCESSION	BD270355		
VERSION	BD270355.1	GI:33080123	
KEYWORDS	JP 2002539775-A/44.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1. (bases 1 to 1464)		
JOURNAL	Rosen, C.A., Ruben, S.M. and Komatsoulis, G.		
TITLE	50 human secreted proteins		
COMMENT	Patent: JP 2002539775-A 44 26-NOV-2002;		
	HUMAN GENOME SCIENCES INC		
	OS Homo sapiens (human)		
	PN JP 2002539775-A/44		
	PD 26-NOV-2002		
	PR 09-MAR-2000 JP 2000606612		
	PF 19-MAR-1999 US 60/125360, 11-JUN-1999 US 60/338626 PR		
	O3-DEC-1999 US 60/168662		
	P1 CRAIG A ROSEN, STEVEN M RUBEN, GEORGE KOMATSOUKIS PC		
	C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P1/00, A61P1/14, PC		
	A61P1/16,		
	PC		
	A61P3/00, A61P3/06, A61P3/10, A61P7/00, A61P7/02, A61P7/06, A61P9/00, PC		
	A61P9/06,		
	PC A61P9/10, A61P9/10, A61P15/00, A61P17/02, A61P17/06, A61P19/00, PC		
	A61P25/00,		
	PC A61P25/14, A61P25/16, A61P25/24, A61P25/28, A61P27/00, A61P27/06,		
	PC A61P29/00,		
	PC A61P31/00, A61P31/04, A61P31/14, A61P31/16, A61P31/18, A61P31/20,		
	PC A61P31/22,		
	PC A61P33/00, A61P33/02, A61P33/06, A61P35/00, A61P35/02, A61P37/00,		

	PC	A61P37/06,	
	PC	A61P37/08,A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC C12N1/21,	
	PC	C12N5/10,C12P21/02,C12Q1/02,C12Q1/68,G01N33/68/(C12P21/02, PC C12R1:19),	
	PC	C12N15/00,A61K37/02,C12N5/00	
	CC	50 human secreted proteins	
	FH	Key Location/Qualifiers	
	FT	source i .1464 /organism='Homo sapiens (human)'	
	FEATURES	Location/Qualifiers	
	source	1..1464 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"	
	ORIGIN		
	Query Match	43.1%; Score 938.6; DB 6; Length 1464;	
	Best Local Similarity	95.7%; Pred. No. 1.2e-198; Indels 8; Gaps 7;	
	Matches 1040; Conservative	0; Mismatches 39;	
Qy	881	AAGTCTCGACGCTTCAGAGCGGGAGCCCCAGCAGGCGCTTGAGCAGAGCTTCGGTGAACGTG	940
Dd	111	AGGCGCCGGGACTTCAGAGCGGGAGCCAGCAGGSCCTTGAGCAGAGCTTCGGCTGAACGTG	170
Qy	941	TGGACCCAGACTCCGAGGCCACCCCAGAGCGCGTGGCACTGTHGTTCAAACCTCATCCA	1000
Dd	171	TGGACCCAGACTCCGAGGCCACCCCAGAGCGCGTGGCACTGTHGTTCAAACCTCATCCA	230
Qy	1001	CTTGGAGCCCTTGGAGTGTCTGTCCTCCCGCCAAACCCTGTTTCCATATGCGCGAGACCCATA	1060
Dd	231	CTTGGAGCCCTTGGAGTGTCTGT - CCCCCCAACCTTGGTTCATATGCCGAGACCCATA	289
Qy	1061	GCCGCCTGCAAGGCAGAGAGACACAGGAGAGCCAGCCCTGAGTCCCGACCTTGGGTGGC	1120
Dd	290	GCGCCCTGCAAGGCAGAGAGACACAGGAGAGCCACCCCTGAGTCCCGACCTTGGGTGGC	349
Qy	1121	GGGGCTGGGTCTCTGCTCCACCCGAGGGCACAGACCGGCTTGCTTTGGCAGGCTGG	1180
Dd	350	GGGGCTTGGGTCTCTGCTCCACCCGAGGGCACAGACCGGCTTGCTTTGGCAGGCTGG	409
Qy	1181	GCCTCTGTGTACCCACTCTCTGGGTGGGTGCAGACCCCTTCCCTTCCACCCCGAGGCTTT	1240
Dd	410	GCCTCTGTGTACCCACTCTCTGGGTGGGTGCAGACCCCTTCCCTTCCACCCCGAGGCTTT	469
Qy	1241	CCAAAGTCTGCTTCTCAGTTTTCAAATGGAACCACTCACTTCCGACGACCCCGACTT	1300
Dd	470	CCAAGCTCTGCTTCTCAGTTTCAAATGGAACCACTCACTTCCGACGACCCCGACTT	529
Qy	1301	ACGAGGAGCATGCCCTCCCTCTGCCCTCATCAAAGCCACAGACCCGGACTCCCTTTCT	1360
Dd	530	ACGAGGAGCATGCCCTCCCTCTGCCCTCATCAAAGCCACAGACCCGGACTCCCTTTCT	589
Qy	1361	GCAACCCAGGCTGGTCCGGCCCCAGAGTGGGGTCCGGCTCTCTCCACTCCGAGGCTCC	1420
Dd	590	GCAACCCAGGCTGGTCCGGCCCCAGAGT - GGGTCCGCTCTCTCACTCCGAGGCTCC	648
Qy	1421	CGGCCCCAAGTGAAGGGGGCCCTGCGGAGGCTCAGACACATGAGATTCAAGGG - TGGG	1478
Dd	649	CGGCCCCAAGTGAAGGGGGCCCTGCGGAGGCTCAGACACATCAGATTCAAGGGCTGGG	708
Qy	1479	GGGGCTTTGCACATACCTGTCCCTTGCTATGAGAGGCTTTTGGGGGCCCTTCCGGGC	1538
Dd	709	GGGGCTTTGGCCACATACCTGTCCCTTGCTATGAGAGGCTTTTGGGGGCCCTTCCGGGC	768
Qy	1539	AGCCCCGGGGCCAGGTAGGGTCTGGGGGCTTAGAGGCTGGGATGGCTCTTGCCCCCAC	1598
Dd	769	AGCCCCGGGGCCAGGTAGGGTCTGGGGGCTTAGAGGCTGGGATGGCTCTTGCCCCCAC	827
Qy	1599	CGCCAGGGGGCAAGCGAGCGCGGGCTGGAGAGCGGGCGGGCTGGGGCTGGGGGGT	1658
Dd	828	CGCCAGGGGGC - ACGCAGAGCGGGCTGGAGAGCGGGCGGGCGGGCTTCGGGGCTGGGGGGT	886

BT008191	BT008191.1	GI:30585220	FLI_CDNA.	synthetic construct	
VERSION	FLI_CDNA.				
KEYWORDS	synthetic construct				
SOURCE	synthetic construct				
ORGANISM	synthetic construct				
REFERENCE	1 (bases 1 to 747)				
AUTHORS	Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M. and Farmer,A.				
TITLE	Cloning of human full-length CDSs in BD Creator(TM) System Donor vector				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 747)				
AUTHORS	Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M. and Farmer,A.				
TITLE	Direct Submission				
JOURNAL	Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA				
COMMENT	This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame.				
FEATURES	Clone distribution: http://bioinfo.clontech.com/orfclones.				
source	1..747				
	/organism="synthetic construct"				
	/mol_type="mRNA"				
	/db_xref="taxon:32630"				
	/clone="GH0750L1.0"				
	/clone_lib="BD Creator(TM) CDS Library derived from MGC collection"				
	/lab_host="DH5alpha T1 resistant"				
	/note="Vector: pDNR-Dual"				
CDS	1..>747				
	/note="Mutations: 746:Stop->Leu"				
	/codon_start=1				
	/transl_table=11				
	/product="Homo sapiens secreted and transmembrane 1"				
	/protein_id="AAP36883.1"				
	/db_xref="GI:30585221"				
	/translation="MOTCPPLAFPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVUS VSWGENTVMSQINISAFSHVNIKLRAHQESAI FNEVAPGYFSDRQWLQVGGVAVQL VIKGARDSHAGLYMHLVGHQNNRQVTLVSGAPQSDPTGFVPVAVVAVFILL VALVNFAYRCSCSQOREKKFFLELPQMKFALRAGAQQLSRASELWTPDSEPTP RPLALVFPSPFLGALLELLSPQLFPFYADPL"				
ORIGIN					
Query Match	33.6%;	Score	732.8;	DB 12;	Length 747;
Best Local Similarity	99.6%;	Pred. No.	6.3e-145;		
Matches	745;	Conservative	0;	Mismatches	2;
	Indels	1;	Gaps	1;	
Qy	314	ATGCAGACCTGCCCCCTGGCATTCCTGGCCACGTTTCCAGGCCCTTGGGACCCCTCTG 373			
Db	1	ATGCAGACCTGCCCCCTGGCATTCCTGGCCACGTTTCCAGGCCCTTGGGACCCCTCTG 60			
Qy	374	TTTTTGGCTGCTCTTGGCTGCTCAGATGAAGCTGGGACAGCCCATCTTGCACAGAG 433			
Db	61	TTTTTGGCTGCTCTTGGCTGCTCAGATGAAGCTGGGACAGCCCATCTTGCACAGAG 120			
Qy	434	GGGCTAGTCTCTGTGTTTGGGGCGAGAACACCGTCATGTCCTGCACATCTTCCACGCC 493			
Db	121	GGGCTAGTCTCTGTGTTTGGGGCGAGAACACCGTCATGTCCTGCACATCTTCCACGCC 180			
Qy	494	TTCTCCCATGTCAACATCAAGCTGCTGCCCGGACGAGAGCGCCCATCTTCAATGAG 553			
Db	181	TTCTCCCATGTCAACATCAAGCTGCTGCCCGGACGAGAGCGCCCATCTTCAATGAG 240			

Db	1	TTCTTGGGGCTCCGGGGCGGAGAGCTGCATCCAGAGAGCGCTCCAGGAGCGGAC 60			
Qy	259	CCGGAGAGTGTTCAGAGCCAGTGAAGGACACGAGGCGCCCAAGTCCACAGCCATGCA 318			
Db	61	CCGGAGAGTGTTCAGAGCCAGTGAAGGACACGAGGCGCCCAAGTCCACAGCCATGCA 120			
Qy	319	GACCTGCCCCCTGGCATTCCTTGGCCACGTTTCCAGGCCCTTGGGACCCCTCTGTTTT 378			
Db	121	GACCTGCCCCCTGGCATTCCTTGGCCACGTTTCCAGGCCCTTGGGACCCCTCTGTTTT 180			
Qy	379	GGCTGCTCTCTTGAAGTGTCAAGTGAAGGCTGGAGACAGCCCATCTGCACAGAGGGGT 438			
Db	181	GGCTGCTCTCTTGAAGTGTCAAGTGAAGGCTGGAGACAGCCCATCTGCACAGAGGGGT 240			
Qy	439	AGTCTCTGTGTTTGGGGCGAGAACACCGTCAATGCTTCTGCAACATCTCCACGCCCTTCTC 498			
Db	241	AGTCTCTGTGTTTGGGGCGAGAACACCGTCAATGCTTCTGCAACATCTCCACGCCCTTCTC 300			
Qy	499	CCATGTCAACATCAAGTGTGTCGCCACGGGAGGAGCGCCATCTTCAATGAGGTGGC 558			
Db	301	CCATGTCAACATCAAGTGTGTCGCCACGGGAGGAGCGCCATCTTCAATGAGGTGGC 360			
Qy	559	TCAGGCTACTTCTCCGGGACGGCTGGCAGCTCCAGGTTTCAGGAGGCGTGGCAGCT 618			
Db	361	TCAGGCTACTTCTCCGGGACGGCTGGCAGCTCCAGGTTTCAGGAGGCGTGGCAGCT 420			
Qy	619	GGTGATCAAGAGCGCGCGGACTCCCATGCTGGGCTGTACATGTGGCACCTCGTGGACA 678			
Db	421	GGTGATCAAGAGCGCGCGGACTCCCATGCTGGGCTGTACATGTGGCACCTCGTGGACA 480			
Qy	679	CCAGAGAAATACAGACAAGTCAAGTGGAGGTTTCAAGTGCAGAACCCACCTCGCCCC 738			
Db	481	CCAGAGAAATACAGACAAGTCAAGTGGAGGTTTCAAGTGCAGAACCCACCTCGCCCC 540			
Qy	739	TGACACTGGGTTCTGGGCTGTGCGAGCGGTGCTCACTGTCTTCACTCTTGTGTGCG 798			
Db	541	CGACACTGGGTTCTGGGCTGTGCGAGCGGTGCTCACTGTCTTCACTCTTGTGTGCG 600			
Qy	799	TCTGGTCAATGTTCCGCTGGTACAGGTGCCGCTGTTCAGCAACGCCGGGAGAGAGTT 858			
Db	601	TCTGGTCAATGTTCCGCTGGTACAGGTGCCGCTGTTCAGCAACGCCGGGAGAGAGTT 660			
Qy	859	CTTCTCTAGAACCCCATAGATGAGTTCGAGCGCTCAGAGCGGAGCCACAGAGGCGCT 918			
Db	661	CTTCTCTAGAACCCCATAGATGAGTTCGAGCGCTCAGAGCGGAGCCACAGAGGCGCT 720			
Qy	919	GAGCAGAGCCTCCGCTGAATGTGGACCCCAAGTCCGAGCCCAACCCAGGCGGTGGC 978			
Db	721	GAGCAGAGCCTCCGCTGAATGTGGACCCCAAGTCCGAGCCCAACCCAGGCGGTGGC 780			
Qy	979	ACTGGTGTCAAAACCTCACACTTGGAGCCCTGGAGCTGCTGCCCCCAACCCCTTGT 1038			
Db	781	ACTGGTGTCAAAACCTCACACTTGGAGCCCTGGAGCTGCTGCCCCCAACCCCTTGT 839			
Qy	1039	TTCCATATGCCGACACCATATAGCGCTGCAAGGACAGAGGAGGACAGAGAGCCAGCC 1098			
Db	840	TTCCATATGCCGACACCATATAGCGCTGCAAGGACAGAGGAGGACAGAGAGTACACC 899			
Qy	1099	CTGAGTGCAGACCTTGGGTGGC-GGGGGCTGGGTCTCTGCTCCACCGGAGGACAGA 1157			
Db	900	CTGAGTGCAGACCTTGGGTGGGCGGGCTGGGTCTCTGCTCCACCGGAGGACAGA 959			
Qy	1158	CACCGGCTGTGTTGGAGGCTGGGCTCTGTGTAC 1193			
Db	960	AGACACCGGGCTTGTGTTGGCAAGGCTTGGGGCTTC 995			

RESULT 14
BT008191
LOCUS
DEFINITION
Synthetic construct Homo sapiens secreted and transmembrane 1 mRNA, partial cds.

```
QY 554 GTGGCTCCAGGTACTTCTCCCGGACCGCTGCGAGCTCCAGTTCAGGAGCGCTGGCA 613
Db 241 GTGGCTCCAGGTACTTCTCCCGGACCGCTGCGAGCTCCAGTTCAGGAGCGCTGGCA 300
QY 614 CAGCTGGTGCATCAAAAGGCGCCCGGACCTCCCATGCTGGGCTATCATGTGGCACTCTGTG 673
Db 301 CAGCTGGTGCATCAAAAGGCGCCCGGACCTCCCATGCTGGGCTATCATGTGGCACTCTGTG 360
QY 674 GGACACCCAGAGAAATAACAGACAAGTCACGCTGGAGTTTCAAGTGCAGAACCCCACTCC 733
Db 361 GGACACCCAGAGAAATAACAGACAAGTCACGCTGGAGTTTCAAGTGCAGAACCCCACTCC 420
QY 734 GGCCTGACACTGGGTTCTGGCTGTGCGAGCGGTGGTCACTGCTGTCTTCACTCTCTTG 793
Db 421 GGCCTGACACTGGGTTCTGGCTGTGCGAGCGGTGGTCACTGCTGTCTTCACTCTCTTG 480
QY 794 GTCGCTCTGGTATGTCCTGTTGCTACAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 853
Db 481 GTCGCTCTGGTATGTCCTGTTGCTACAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 540
QY 854 AAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 913
Db 541 AAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
QY 914 GGCCTGACAGAGCTCTCGCTGAAGTGTGGACCCAGAGTCCGAGCTCCGAGCCCAAGGCG 973
Db 601 GGCCTGACAGAGCTCTCGCTGAAGTGTGGACCCAGAGTCCGAGCTCCGAGCCCAAGGCG 660
QY 974 CTGGCACTGGTGTTCACACCTTCAAGTGTGGACCTTCAAGTGTGGACCTTCAAGTGT 1033
Db 661 CTGGCACTGGTGTTCACACCTTCAAGTGTGGACCTTCAAGTGTGGACCTTCAAGTGT 719
QY 1034 CTGTTTCTCATATGCGCGAGACCCATAG 1061
Db 720 CTGTTTCTCATATGCGCGAGACCCATAG 747

RESULT 15
AC084737 166458 bp DNA linear HTG 11-NOV-2000
LOCUS Homo sapiens chromosome UL clone RP13-516M14, WORKING DRAFT
DEFINITION AC084737
ACCESSION AC084737
VERSION AC084737.2 GI:11138191
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 166458)
Waterston,R.H.
Direct Submission
Submitted (10-NOV-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Nov 11, 2000 this sequence version replaced gi:11138191.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.FH0516M14
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
```

```
Assembly program: Phrap; version 0.990319
Consensus quality: 147705 bases at least Q40
Consensus quality: 155070 bases at least Q30
Consensus quality: 158378 bases at least Q20
Insert size: 172000; agarose-fp
Insert size: 164058; sum-of-contigs
Quality coverage: 3.66 in Q20 bases; agarose-fp
Quality coverage: 3.90 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
1192: contig of 1192 bp in length
1292: gap of unknown length
1293
1293
2830: contig of 1538 bp in length
2831
2930: gap of unknown length
2931
4485: contig of 1555 bp in length
4486
4585: gap of unknown length
7134: contig of 2549 bp in length
7135
7235
9400: contig of 2166 bp in length
9401
11389: contig of 1889 bp in length
11390
11489: gap of unknown length
11490
13890: contig of 2401 bp in length
13891
13990: gap of unknown length
13991
16786: contig of 2796 bp in length
16787
16886: gap of unknown length
16887
20616: contig of 3730 bp in length
20617
20716: gap of unknown length
20717
23530: contig of 2814 bp in length
23531
26912: contig of 3282 bp in length
26913
27012: gap of unknown length
27013
29489: contig of 2477 bp in length
29490
29589: gap of unknown length
29590
32528: contig of 2939 bp in length
32529
32628: gap of unknown length
32629
38475: contig of 5847 bp in length
38476
38575: gap of unknown length
38576
43098: contig of 4523 bp in length
43099
43198: gap of unknown length
43199
46803: contig of 3605 bp in length
46804
46903: gap of unknown length
46904
51761: contig of 4858 bp in length
51762
51861: gap of unknown length
51862
57363: contig of 5502 bp in length
57364
57463: gap of unknown length
57464
65221: contig of 7758 bp in length
65222
65321: gap of unknown length
65322
72371: contig of 7050 bp in length
72372
72471: gap of unknown length
72472
80691: contig of 8220 bp in length
80692
80791: gap of unknown length
80792
86118: contig of 7827 bp in length
86119
88718: gap of unknown length
88719
109660: contig of 20942 bp in length
109661
109760: gap of unknown length
109761
132756: contig of 22996 bp in length
132757
132856: gap of unknown length
132857
164458: contig of 33602 bp in length.
Location/Qualifiers
1..166458
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="UL"
/clone="RP13-516M14"
FEATURES
source
```

Search completed: March 20, 2005, 04:49:43
Job time : 9545 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 23:11:41 ; Search time 1185 Seconds
(without alignments)
10890.326 Million cell updates/sec

Title: US-10-080-522-2
Perfect score: 2180
Sequence: 1 ATTCTGCTTCTTTAGCGT.....TTAAATAAGGTTTCGTCGC 2180

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2168.8	99.5	2180	2	AAV32446 K12 CDNA
2	1951	89.5	2000	4	AAC88152 Human K12
3	1951	89.5	2000	6	ABL65927 Lung canc
4	1951	89.5	2000	6	ABK84431 Human CDN
5	1951	89.5	2000	12	ADJ74845 Marker ge
6	1951	89.5	2000	13	ADR52746 Drug ther
7	1951	89.5	2000	13	ADP24663 PRO polyp
8	1751	80.3	1985	12	ADQ86642 Human tum
9	1751	80.3	1985	13	ADQ85525 Human tum
10	938.6	43.1	1464	3	AAC59772 Human sec
11	912.8	41.9	1064	2	AAX82081 Human SIG
12	692.4	31.8	958	10	ADB36352 Human imm
13	673.8	30.9	942	10	ADB36353 Human imm
14	636.4	29.2	930	9	AAL60549 Human org
15	512.8	23.5	869	9	AAL60550 Human org
16	506.2	23.2	592	10	ADF81837 Leukaemia
17	392.8	18.0	763	9	AAL60548 Human org
18	363.8	16.7	401	10	ADF80286 Leukaemia
19	339.4	15.6	765	9	AAL60551 Human org
20	222	10.2	3026	12	ADM18621 Human chr

c

ALIGNMENTS

RESULT 1
AAV32446
ID AAV32446 standard; cDNA; 2180 BP.
XX AC AAV32446;
XX DT 26-OCT-1998 (first entry)
XX DE K12 CDNA sequence.
XX KW DNA probe; CD7 HSI DNase hypersensitive site; mRNA northern blot;
KW human erythroleukemic; HEL; K562 cell line; clone; breast cancer;
KW ovarian cancer; malignant; tumour; ss.
XX OS Homo sapiens.
XX FH Key
XX CDS Location/Qualifiers
FT CDS 314..1061
FT /*tag= a
FT /transl_except= (pos:1028..1031, aa:Gln)
FT /note= "this codon has an apparent 1 nucleotide
FT insertion, which alters the reading frame"
XX PN WO9822502-A1.
XX PD 28-MAY-1998.
XX PP 24-NOV-1997; 97WO-US021517.
XX PR 22-NOV-1996; 96US-00755559.
XX PA (UYDU-) UNIV DUKE.
XX PI Kaufman RE, Slentz-Kesler KA;
XX DR WPI; 1998-312415/27.
XX DP P-PSDB; AAW48811.
XX PT New isolated K12 protein gene - which is over expressed in certain
XX neoplastic cells, used to develop products for tumour detection and
XX treatment.
XX PS Claim 8; Fig 2A-2B; 44pp; English.

21 197 9.0 212 2 AAT26356 Aat26356 Human gen
22 157 7.2 1076 12 ADJ75761 Adj75761 Marker ge
23 155.4 7.1 1112 4 AAC88154 Aac88154 Mouse K12
24 155.4 7.1 1112 12 ADJ75760 Adj75760 Marker ge
25 74.8 3.4 79 4 AAC89734 Aac89734 Human gas
26 70 3.2 90 4 AAC89735 Aac89735 Human gas
27 60 2.8 60 6 ABN38487 Abn38487 Human spl
28 56.8 2.6 1127 3 AAA02477 Aaa02477 Human col
29 56 2.6 2000 8 ADA71938 Ada71938 Rice gene
30 54 2.5 100301 6 ABQ88176 Abq88176 Human ost
31 54 2.5 100301 12 ADQ20754 Adq20754 Human sof
32 53.2 2.4 114955 2 AAX53491 Aax53491 Human ade
33 51.4 2.4 2000 8 ADJ78312 Adj78312 Human per
34 51.2 2.3 5452 10 ADC86736 Adc86736 Human gpc
35 51 2.3 114955 2 AAX53491 Aax53491 Human ade
36 49.8 2.3 660 5 AAS68601 Aas68601 DNA encod
37 49.6 2.3 1679 8 ABZ78139 Abz78139 Human can
38 49.4 2.3 4667 12 ADQ22939 Adq22939 Human sof
39 49.4 2.3 5452 10 ADC86736 Adc86736 Human gpc
40 48.4 2.2 3897 12 ADJ78312 Adj78312 Human per
41 48.4 2.2 22210 12 ADJ78311 Adj78311 Human per
42 48 2.2 647 4 AAH06718 Aah06718 Human CDN
43 48 2.2 2437 4 AAH18557 Aah18557 Human CDN
44 47.6 2.2 8459 4 AAC89557 Aac89557 Human his
45 47.6 2.2 8459 8 ACC47997 Acc47997 Human his

xx This present sequence represents the K12 cDNA, which has been located on
cc chromosome 17q25. To obtain the cDNA sequence a 500 bp DNA probe, which
cc can be located just upstream of the CD7 Hsi DNase hypersensitive site,
cc was used against a mRNA northern blot. From this a 1.8kb transcript was
cc detected in the human erythroleukemic cell line HEL. The probe was then
cc used to screen a human erythroleukemic cell line K562 cDNA library, from
cc which several clones were identified and isolated that constituted a
cc 1.8kb cDNA. This cDNA was designated K12 and was found to have a single
cc open reading frame as well as being in the same orientation as CD7. The
cc K12 gene was found to be expressed in both breast and ovarian cancer
cc cells at a much higher level than any other malignant or normal tissue
cc that was examined, thus enabling the K12 to be a useful protein in tumour
cc detection and treatment

XX
CC
COMM-220 3180 BP. 381 A. 730 C. 639 G: 420 T: 0 U: 0 Other;

Query Match	99.5%	Score 2168.8	DB 2	Length 2180
Best Local Similarity	99.7%	Pred. NO. 0		
Matches 2173	Conservative	0	Mismatches 7	Indels 0
				Gaps 0

QY	1	ATTCTGCTTCTTTAGCGTGAACGCGGTGCGTCCCGTGAATAATAAATTCAC	60
DB	1	ATTCTGCTTCTTTAGCGTGAACGCGGTGCGTCCCGTGAATAATAAATTCAC	60
QY	61	CGTCAACGCTTGTTGTGAACGCGGTGTTCCGAAACCTTGAGAGCTTCCCGTAACCCAG	120
DB	61	CGTCAACGCTTGTTGTGAACGCGGTGTTCCGAAACCTTGAGAGCTTCCCGTAACCCAG	120
QY	121	CTCCTTCTCATCTGGGAGTGGGTCCGCGCGGTCCGCGCTTCCCTCCCTGGCGCCCTC	180
DB	121	CTCCTTCTCATCTGGGAGTGGGTCCGCGCGGTCCGCGCTTCCCTCCCTGGCGCCCTC	180
QY	181	CCCTCTCGTGTCTTTTCATTTTCTTGGGGTTCGGGGCGCGGAGAGCTGCATCCACAGGA	240
DB	181	CCCTCTCGTGTCTTTTCATTTTCTTGGGGTTCGGGGCGCGGAGAGCTGCATCCACAGGA	240
QY	241	GGCGGTCCAGGAGCGGACCCGGGAGTGTTCAGAGCCAGTGAACAAGACACCAAGGGCCCCA	300
DB	241	GGCGGTCCAGGAGCGGACCCGGGAGTGTTCAGAGCCAGTGAACAAGACACCAAGGGCCCCA	300
QY	301	AGTCCACACAGGCATCGACACCTGCCCTCGCATTCCTGTGCCACAGTTTCCAGGCCCT	360
DB	301	AGTCCACACAGGCATCGACACCTGCCCTCGCATTCCTGTGCCACAGTTTCCAGGCCCT	360
QY	361	TGGGACCCCTCTGTTTGGCTGCCTCTTGTAGTGTCTCAGATGAAGCTGGGACAGGCC	420
DB	361	TGGGACCCCTCTGTTTGGCTGCCTCTTGTAGTGTCTCAGATGAAGCTGGGACAGGCC	420
QY	421	CATCTGCACAGAGGGGTAGTCTCTGTCTTTGGGCGAGAAACA CGTCAATGCTCTGCA	480
DB	421	CATCTGCACAGAGGGGTAGTCTCTGTCTTTGGGCGAGAAACA CGTCAATGCTCTGCA	480
QY	481	CATCTCAACGCCCTTCTCCCATGTCAACATCAGCTGCGTGCCACCGGCGAGAGCGC	540
DB	481	CATCTCAACGCCCTTCTCCCATGTCAACATCAGCTGCGTGCCACCGGCGAGAGCGC	540
QY	541	CATCTTCAATAGTGCTCCAGGCTACTTCTCCGGGACGGGTGCGAGCTCCAGGTTCA	600
DB	541	CATCTTCAATAGTGCTCCAGGCTACTTCTCCGGGACGGGTGCGAGCTCCAGGTTCA	600
QY	601	GGGAGGCGTGGGACAGCTGGTGATCAAAAGGCGCGGGACTCCCATGCTGGGTGTACAT	660
DB	601	GGGAGGCGTGGGACAGCTGGTGATCAAAAGGCGCGCGGACTCCCATGCTGGGTGTACAT	660
QY	661	GTGCACCTCGTGGGACACAGAGAAATAACAGAAAGTCAAGCTGGAGGTTTCAGGTGC	720
DB	661	GTGCACCTCGTGGGACACAGAGAAATAACAGAAAGTCAAGCTGGAGGTTTCAGGTGC	720
QY	721	AGAACCCCAAGTCCGCCCTTGACATCTGGGTTCTGGCTGTGCCAGCGGTGTCACTGCTGT	780
DB	721	AGAACCCCAAGTCCGCCCTTGACATCTGGGTTCTGGCTGTGCCAGCGGTGTCACTGCTGT	780

Db 1861 TTTGAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAACAAACCGACTCACAACCC 1920
Qy 1921 TCCAAAGTAGCTCAAAATGCAATTTTAAATGGAATAACAAAATCTGAAAGAAAGCTCTT 1980
Db 1921 TCCAAAGTAGCTCAAAATGCAATTTTAAATGGAATAACAAAATCTGAAAGAAAGCTCTT 1980
Qy 1981 TAGTGGCTTTAAGCCCAAAACGCTCCCTAAGGCGTCTCGAGATGAAGACGGGGGGAGC 2040
Db 1981 TAGTGGCTTTAAGCCCAAAACGCTCCCTAAGGCGTCTCGAGATGAAGACGGGGGGAGC 2040
Qy 2041 CCCAGCCAGGTGAGACCCCGCAGACGGCGCGCCGCTGACCGAGGCTCGCACAG 2100
Db 2041 CCCAGCCAGGTGAGACCCCGCAGACGGCGCGCCGCTGACCGAGGCTCGCACAG 2100
Qy 2101 CCGGCGCGCTCAGGCGTGGGCGGAGCCAGGCTCAAGAGGGCGCGTTTGTGTCTCGGG 2160
Db 2101 CCGGCGCGCTCAGGCGTGGGCGGAGCCAGGCTCAAGAGGGCGCGTTTGTGTCTCGGG 2160
Qy 2161 TTAATAAAGTTCCGTCG 2180
Db 2161 TTAATAAAGTTCCGTCG 2180

RESULT 2

ID AAC88152 standard; cDNA; 2000 BP.
AC AAC88152;
XX
DT 13-MAR-2001 (first entry)
DE Human K12 nucleotide sequence SEQ ID NO:3.
XX
KW Human; CD7; K12; cognate ligand; cluster of differentiation; cancer;
KW identification; inhibiting T cell proliferation; HIV; infection;
KW activating natural killer cell proliferation; leukaemia; lymphoma;
KW sepsis; graft versus host disease; autoimmune disease; arthritis;
KW multiple sclerosis; rheumatoid arthritis; psoriatic arthritis; lupus;
KW scleroderma; psoriasis; atopic dermatitis; type I diabetes mellitus;
KW Hashimoto's thyroiditis; pernicious anaemia; Addison's disease; uveitis;
KW myasthenia gravis; psoriasis; Guillain-Barre Syndrome; Grave's disease;
KW systemic lupus erythematosus; dermatomyositis; asthma; eczema;
KW atypical dermatitis; contact dermatitis; eczematous dermatitis;
KW seborrheic dermatitis; rhinitis; ss.
XX
OS Homo sapiens.
XX
PN WO200073333-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US014612.
XX
PR 28-MAY-1999; 99US-0136450P.
XX
PA (INMV) IMMUNEX CORP.
XX
PI Lyman SD, Fanslow WC;
XX
DR WPI; 2001-061511/07.
XX
DR P-PSDB; AAB36658.
XX
PT Stimulating intracellular signaling of CD7 comprises contacting a cell
PT expressing CD7 with recombinant K12 protein, the cognate ligand of CD7,
PT to inhibit T cell proliferation and/or activate natural killer cell
PT proliferation.
XX
PS Disclosure; Page 36-38; 42pp; English.
XX
CC The present invention describes a method for stimulating (S) the
CC intracellular signalling of CD (cluster of differentiation) 7 comprising
CC contacting a cell that expresses CD7 with a recombinant K12 protein (I),

CC the cognate ligand of CD7. (S) is useful for inhibiting T cell
CC proliferation and/or activating NK (natural killer) cell proliferation
CC and/or inducing NK toxicity in a mammal which involves administration of
CC K12 protein. It is also used for treating HIV-1 infection, cancer (T cell
CC leukaemia, acute lymphocytic leukaemia, cutaneous T cell lymphoma),
CC bacterial and viral infections, mediated by CD7. In the case of treating
CC T cell leukaemia the soluble K12 protein is covalently attached to a
CC toxin. A disease mediated by CD7 such as sepsis, graft versus host
CC disease due to transplantation, autoimmune diseases, multiple sclerosis,
CC arthritis, rheumatoid arthritis, psoriatic arthritis, scleroderma, lupus,
CC psoriasis, atopic dermatitis, type I diabetes mellitus, Hashimoto's
CC thyroiditis, pernicious anaemia, Addison's disease, myasthenia gravis,
CC uveitis, psoriasis, Guillain-Barre Syndrome, Grave's disease, systemic
CC lupus erythematosus and dermatomyositis, asthma, eczema, atypical
CC dermatitis, contact dermatitis, other eczematous dermatitis, seborrheic
CC dermatitis, and rhinitis is also treated by administering a K12
CC antagonist (neuralising antibody). The present sequence encodes the
CC human K12 protein, which is given in the exemplification of the present
CC invention
XX
SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;
Query Match 89.5%; Score 1951; DB 4; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
Qy 196 ATTTTCTGGGGCTCCGGGGCGCGAGAAAGCTGCATCCAGAGGAGCGCTCCAGGAGCG 255
Db 1 ATTTTCTGGGGCTCCGGGGCGCGAGAAAGCTGCATCCAGAGGAGCGCTCCAGGAGCG 60
Qy 256 GACCCGGGAGTGTTCAGAGCCAGTGAACAGGACAGGCGGCCAACAGCCAT 315
Db 61 GACCCGGGAGTGTTCAGAGCCAGTGAACAGGACAGGCGGCCAACAGCCAT 120
Qy 316 GCAGACCTGCCCTGGCATTCCTGGCCAGCTTTCCAGGAGCGCTTCCTGTT 375
Db 121 GCAGACCTGCCCTGGCATTCCTGGCCAGCTTTCCAGGAGCGCTTCCTGTT 180
Qy 376 TTTGGCTGCTCTTGGAGTGTCTGAGATGAAGGCTGGGACAGCCCATCTGCACAGAGGG 435
Db 181 TTTGGCTGCTCTTGGAGTGTCTGAGATGAAGGCTGGGACAGCCCATCTGCACAGAGGG 240
Qy 436 GGTAGTCTCTGTGCTTGGGGCGAGAACACCGTCAATGCTGTGCAACATCTCCAAGCCTT 495
Db 241 GGTAGTCTCTGTGCTTGGGGCGAGAACACCGTCAATGCTGTGCAACATCTCCAAGCCTT 300
Qy 496 CTCCCATGTCAACATCAAGCTGCGTCCAGCGGACAGAGCGCCATCTCAATGAGGT 555
Db 301 CTCCCATGTCAACATCAAGCTGCGTCCAGCGGACAGAGCGCCATCTCAATGAGGT 360
Qy 556 GGCTCCAGGCTACTTCTCCCGGACCGCTGGCAGCTCCAGGTTCAAGGAGCGCTGGCACA 615
Db 361 GGCTCCAGGCTACTTCTCCCGGACCGCTGGCAGCTCCAGGTTCAAGGAGCGCTGGCACA 420
Qy 616 GCTGGTGATCAAAAGGCGCCGGGACTCCCATGCTGGGTGTACATGTGGCACCTCTGGGG 675
Db 421 GCTGGTGATCAAAAGGCGCCGGGACTCCCATGCTGGGTGTACATGTGGCACCTCTGGGG 480
Qy 676 ACACGAGAGAAATAACAGACAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 735
Db 481 ACACGAGAGAAATAACAGACAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 540
Qy 736 CCCTGACACTGGGTCTGGGCTGTGCCAGCGGTGTCACTGTCTTCATCTCTCTGTT 795
Db 541 CCCTGACACTGGGTCTGGGCTGTGCCAGCGGTGTCACTGTCTTCATCTCTCTGTT 600
Qy 796 CGCTCTGTGATGTTGCGCTGGTACAGGTGCGCTGTTCCAGCAACCGCCGGGAGAGAA 855
Db 601 CGCTCTGTGATGTTGCGCTGGTACAGGTGCGCTGTTCCAGCAACCGCCGGGAGAGAA 660
Qy 856 GTTCTTCTCTAGAACCCAGATGAAGTTCAGAGCGGCTTCAGAGCGGAGCGCCAGCAGG 915
Db 661 GTTCTTCTCTAGAACCCAGATGAAGTTCAGAGCGGCTTCAGAGCGGAGCGCCAGCAGG 720

1996 CCAAACAGCTCCCTAAGCGCTCTCTGAGATGAAGACGGGGGGAG-CCCAGCCAGGTGGA 2054
1800 CCAAACAGCTCCCTAAGCGCTCTCTGAGATGAAGACGGGGGGAGCCCCCAGCCAGGTGGA 1859
2055 GACCCCGCAGGACGCGCGCGCGCGCGCTGACCCGAGGCTTCGACAGCGCGCGCGCGCTGAG 2114
1860 GACCCCGCAGGACGCGCGCGCGCGCGCTGACCCGAGGCTTCGACAGCGCGCGCGCGCTGAG 1919
2115 GGTTCGGGCC-GAGCCAGGCTTCAAGAGGGCGGCTTGTGTCTCGGGTTAAATAAGGTT 2173
1920 GGTTCGGGCCGAGCCAGGCTTCAAGAGGGCGGCTTGTGTCTCGGGTTAAATAAGGTT 1979
2174 CCGTCCG 2180
1980 CCGTCCG 1986

RESULT 3
ABL65927
ID ABL65927 standard; DNA; 2000 BP.
XX ABL65927;
AC ABL65927;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4264.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
gene; ds.
XX
OS Homo sapiens.
XX
FN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-023133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 22-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.

916 CTTGAGCAGAGCTCCGCTGAATCTGTGACCCCGAGCTCCGAGCCACCCCAAGCGCGCT 975
721 CTTGAGCAGAGCTCCGCTGAATCTGTGACCCCGAGCTCCGAGCCACCCCAAGCGCGCT 780
976 GGCACCTGTGTTCAACCCCTACCACTTGGAGCCCTGGAGCTGTGTCTCCCGCCCAACCCCT 1035
781 GGCACCTGTGTTCAACCCCTACCACTTGGAGCCCTGGAGCTGTGTCTCCCGCCCAACCCCT 839
1036 TGTGTTCCATATGCGCAGACCCATAGCCGCTTGAAGGCAGAGAGGACACAGGAGAGCCA 1095
840 TGTGTTCCATATGCGCAGACCCATAGCCGCTTGAAGGCAGAGAGGACACAGGAGAGCCA 899
1096 GGCCTGAGTGCAGACTTGGGTGGGGGCTGTGGGTCTCTGTCTCCACCCCGAGGGGACA 1155
900 GGCCTGAGTGCAGACTTGGGTGGGGGCTGTGGGTCTCTGTCTCCACCCCGAGGGGACA 959
1156 GACACCGGCTTGTGAGGAGCTGGGCTCTGTGTACCCACTCTCTGGGTGCTGTCAGAC 1215
960 GACACCGGCTTGTGAGGAGCTGGGCTCTGTGTACCCACTCTCTGGGTGCTGTCAGAC 1019
1216 CTTTCCCTCCACCCCGAGGCTTCCAAAGCTCTGCTTCTCAGTTTCCAAATGGAACC 1275
1020 CTTTCCCTCCACCCCGAGGCTTCCAAAGCTCTGCTTCTCAGTTTCCAAATGGAACC 1079
1276 ACTCACTCCGAGACCCGACTTACAGAGGAGCATGCCCTCTGCTCTGCTCATCAA 1335
1080 ACTCACTCCGAGACCCGACTTACAGAGGAGCATGCCCTCTGCTCTGCTCATCAA 1139
1336 ACCACAGACCCGAGCTCCCTTTCTGCCACCCCGAGGCTGTGCGGCCCGAGGTGGGT 1395
1140 ACCACAGACCCGAGCTCCCTTTCTGCCACCCCGAGGCTGTGCGGCCCGAGGTGGGT 1199
1396 CCGCTCTCTCACTCCAGGCTTCCGCGCCCAAGTGAAGGGGCGCTTCCGAGGCTCAG 1455
1200 CCGCTCTCTCACTCCAGGCTTCCGCGCCCAAGTGAAGGGGCGCTTCCGAGGCTCAG 1259
1456 ACACACTGGAGTTGAGGCTGGGGGGCTTGGCACAATACCTCTGCTTGGCTATGACA 1515
1260 ACACACTGGAGTTGAGGCTGGGGGGCTTGGCACAATACCTCTGCTTGGCTATGACA 1319
1516 GGCCTTTCGGGGCCCTTCCGCGGAGCCGCGGGGCGGAGTAGGGTCTGGGGGCTTAGAG 1575
1320 GGCCTTTCGGGGCCCTTCCGCGGAGCCGCGGGGCGGAGTAGGGTCTGGGGGCTTAGAG 1379
1576 GCTGGGATGGTCTCTGCGCCCAACCGCCAGGGGCGAGCGAGCGGGCTGGGAGCGGC 1635
1380 GCTGGGATGGTCTCTGCGCCCAACCGCCAGGGGCGAGCGAGCGGGCTGGGAGCGGC 1439
1636 GCGCGCGCTCGGGCTGGGGGCTCAGGTGGAACGCTGCTCCGGGGCTGTGCGGCATCCC 1695
1440 GCGCGCGCTCGGGCTGGGGGCTCAGGTGGAACGCTGCTCCGGGGCTGTGCGGCATCCC 1499
1696 TCAGTCCCTCGGCCACCCGGGGGCTGCTCCCTGTCGTCACGCGGGGCTTTGATGCTTT 1755
1500 TCAGTCCCTCGGCCACCCGGGGGCTGCTCCCTGTCGTCACGCGGGGCTTTGATGCTTT 1559
1756 TGGACCCAGATCTGTTATGCTTTGCTTCTGCTCACTGCGCGGGGCTTTGATGCTTT 1815
1560 TGGACCCAGATCTGTTATGCTTTGCTTCTGCTCACTGCGCGGGGCTTTGATGCTTT 1619
1816 CATCTGATGGGTGGAAAAATCACCGGGAATCCCGCTTCAAGTTCTTTGAAAAAGTTCCA 1875
1620 CATCTGATGGGTGGAAAAATCACCGGGAATCCCGCTTCAAGTTCTTTGAAAAAGTTCCA 1679
1876 TGACTCGAATATCTGAATGAAGAAACAAACCCGACTCACAACCTCCAGTAGCTCAA 1935
1680 TGACTCGAATATCTGAATGAAGAAACAAACCCGACTCACAACCTCCAGTAGCTCAA 1739
1936 ATGCAATTTTAAATGAAACAAATACTGAAAGAACCTCTTTAGTGGCTTTAAGCC 1995
1740 ATGCAATTTTAAATGAAACAAATACTGAAAGAACCTCTTTAGTGGCTTTAAGCC 1799

PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX Claim 1; SEQ ID NO 4264; 44pp; English.
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
XX tumour
XX
SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 89.5%; Score 1951; DB 6; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 196 ATTTTCTGGGCTCGGGCGCGGAGAGCTGCATCCAGAGAGCGGTCCAGAGCG 255
DB |||||
QY 256 GACCCGGAGTGTTTCAAGAGCCAGTGACAGGACCCAGGGGCCAAGTCCCAACAGCCAT 315
DB 61 GACCCGGAGTGTTTCAAGAGCCAGTGACAGGACCCAGGGGCCAAGTCCCAACAGCCAT 120
QY 316 CGAGACCTGCCCCCTCGCATTCCTGGCCACGTTTCCAGGCGCTTGGAGCCCTCTGT 375
DB 121 CGAGACCTGCCCCCTCGCATTCCTGGCCACGTTTCCAGGCGCTTGGAGCCCTCTGT 180
QY 376 TTGGCTGCTCTTGTAGTGCTCAGATGAAGCTGGAGACGCCCATCTGCACAGAGGG 435
DB 181 TTGGCTGCTCTTGTAGTGCTCAGATGAAGCTGGAGACGCCCATCTGCACAGAGGG 240
QY 436 GGTAGTCTGTGTCTTGGGGGAGAACACCGTCAATGCTCTGCAACATCTCCAAGCCCTT 495
DB 241 GGTAGTCTGTGTCTTGGGGGAGAACACCGTCAATGCTCTGCAACATCTCCAAGCCCTT 300
QY 496 CTCCCATGTCAACATCAAGCTCGTCCCAAGGGGAGGAGCGCCATCTTTCAATGAGGT 555
DB 301 CTCCCATGTCAACATCAAGCTCGTCCCAAGGGGAGGAGCGCCATCTTTCAATGAGGT 360
QY 556 GCCTCCAGGCTACTTCTCCGGGAGCGGTGGAGCTCAGGTTCAGGAGGCGGTGGCACA 615
DB |||||

DB 361 GGCTCCAGGCTACTTCTCCCGGGACGGCTGGCAGCTCCAGGTTCCAGGAGGCGTGGCACA 420
QY 616 GCTGTGATCAAAAGGCGCCGGGACTCCCATGCTGGGTGTACATGTGGACACTCGTGGG 675
DB 421 GCTGTGATCAAAAGGCGCCGGGACTCCCATGCTGGGTGTACATGTGGACACTCGTGGG 480
QY 676 ACACAGAGAAATAACAGACAAAGTCACTCGTGGAGGTTTCAAGTGAGAACCCAGTCCGC 735
DB 481 ACACAGAGAAATAACAGACAAAGTCACTCGTGGAGGTTTCAAGTGAGAACCCAGTCCGC 540
QY 736 CCCTGACACTGGGTTCGGCTGTGCGCAGCGGTGTGCTACTGCTGCTTCATCTCTTGGT 795
DB 541 CCCTGACACTGGGTTCGGCTGTGCGCAGCGGTGTGCTACTGCTGCTTCATCTCTTGGT 600
QY 796 CGCTCTGCTATGTTTCGCTGTGTAAGGTGCGCTGTTTCCAGAGCAACCGCGGAGAGAA 855
DB 601 CGCTCTGCTATGTTTCGCTGTGTAAGGTGCGCTGTTTCCAGAGCAACCGCGGAGAGAA 660
QY 856 GTTCTTCTCTAGAACCCAGATGAAGGTGCGAGCCCTCAGAGCGGAGCCACAGAGGG 915
DB 661 GTTCTTCTCTAGAACCCAGATGAAGGTGCGAGCCCTCAGAGCGGAGCCACAGAGGG 720
QY 916 CCTGAGCAGAGCTCCGCTGAACCTGTGACCCCGAGACTCCGAGCCCAACAGGCGCT 975
DB 721 CCTGAGCAGAGCTCCGCTGAACCTGTGACCCCGAGACTCCGAGCCCAACAGGCGCT 780
QY 976 GGCACCTGTTTCAAAACCCCTCACCACTTGGAGCCCTGGAGCTGTGCTCCCGCCCAACCT 1035
DB 781 GGCACCTGTTTCAAAACCCCTCACCACTTGGAGCCCTGGAGCTGTGCT-CCCCCAACCT 839
QY 1036 TGTTCATATGCGCAGACCCATAGCCCTGCAAGCGCAGAGAGGACACAGGAGAGCA 1095
DB 840 TGTTCATATGCGCAGACCCATAGCCCTGCAAGCGCAGAGAGGACACAGGAGAGCA 899
QY 1096 GCTCTGAGTCGCGACCTTGGGTGGGGGCGCTGGGTCTCTGTCACCCCGGAGGCGACA 1155
DB 900 GCTCTGAGTCGCGACCTTGGGTGGGGGCGCTGGGTCTCTGTCACCCCGGAGGCGACA 959
QY 1156 GACACCGGCTTGTGTCGAGGCTGGGCTCTGTGTACCCACTCTCTGGGTGCTGCAGAC 1215
DB 960 GACACCGGCTTGTGTCGAGGCTGGGCTCTGTGTACCCACTCTCTGGGTGCTGCAGAC 1019
QY 1216 CTTTCCCTCCACCCCGCAGGTCTTCCAAAGCTCTGCTTCTCAGTTTCCAAATGGAACC 1275
DB 1020 CTTTCCCTCCACCCCGCAGGTCTTCCAAAGCTCTGCTTCTCAGTTTCCAAATGGAACC 1079
QY 1276 ACCTCACCTCGCAGCACCAGCTTACAGAGCCGATGCCCTCCCTGCTGCCCTCATCAA 1335
DB 1080 ACCTCACCTCGCAGCACCAGCTTACAGAGCCGATGCCCTCCCTGCTGCCCTCATCAA 1139
QY 1336 ACCCACAGCCCGGACTCCCTTTCTGCGACCCCGAGGCTGGTCCGGCCCGCAGGTGGGGT 1395
DB 1140 ACCCACAGCCCGGACTCCCTTTCTGCGACCCCGAGGCTGGTCCGGCCCGCAGGTGGGGT 1199
QY 1396 CGCTCTCTCACTCCAGGGCTCGCGCCCAAAGTGAAGGGGCGCTCTGCCGAGGCTCAG 1455
DB 1200 CGCTCTCTCACTCCAGGGCTCGCGCCCAAAGTGAAGGGGCGCTCTGCCGAGGCTCAG 1259
QY 1456 ACACACTGGAGTTCAGGGCTGGGGGCGCTTGGGCACATACCTGCTCCCTTGGCTATGACA 1515
DB 1260 ACACACTGGAGTTCAGGGCTGGGGGCGCTTGGGCACATACCTGCTCCCTTGGCTATGACA 1319
QY 1516 GGCTTTGGGGGCGCTTCCGCGCAGACCCCGGGGGCCGAGGTAGGTCTGGGGGCTTAGAG 1575
DB 1320 GGCTTTGGGGGCGCTTCCGCGGCGACGCCCGGGGGCGGAGGTAGGTCTGGGGGCTTAGAG 1379
QY 1576 GCTGGGATGGTCTCTGGGCCCAACCGCAGGGGGCAAGCGCAGGCGGGCTGGGAGCGGC 1635
DB 1380 GCTGGGATGGTCTCTGGGCCCAACCGCAGGGGGCAAGCGCAGGCGGGCTGGGAGCGGC 1439
QY 1636 GCGGCGGCTCGGGCTGGGGGCTCAGGTGAGACGCTGCTCCCGGGGCTGGTCCGCGATCCC 1695
DB 1440 GCGGCGGCTCGGGCTGGGGGCTCAGGTGAGACGCTGCTCCCGGGGCTGGTCCGCGATCCC 1499

QY	1696	TCAGTCCCTCGGCCACCCGGGGTGCCTCCCTGTCGCCACCGACCTGCCAGCGCTCTT	1755
Db	1500	TCAGTCCCTCGGCCACCCGGGGTGCCTCCCTGTCGCCACCGACCTGCCAGCGCTCTT	1559
QY	1756	TGAGCCAGATCTGTTTCATGCTTTTGTCTTCTGTCACCTGGGGGGCTTTGATGCTT	1815
Db	1560	TGAGCCAGATCTGTTTCATGCTTTTGTCTTCTGTCACCTGGGGGGCTTTGATGCTT	1619
QY	1816	CATCTGTATGGGGTGAATAATCACCGGAATCCCTTCAGTTCTTTGAAAAAGTTCCA	1875
Db	1620	CACTGTATGGGGTGAATAATCACCGGAATCCCTTCAGTTCTTTGAAAAAGTTCCA	1679
QY	1876	TGACTCGAATATCTGAATGAAGAAAAACAAACCGACTCACAACCTCAAGTAGTCCAA	1935
Db	1680	TGACTCGAATATCTGAATGAAGAAAAACAAACCGACTCACAACCTCAAGTAGTCCAA	1739
QY	1936	ATGCAATTTTAAATGGAACAAAAATCTGAAGAAACGTCCTTTAGTGGCTTTAAGCC	1995
Db	1740	ATGCAATTTTAAATGGAACAAAAATCTGAAGAAACGTCCTTTAGTGGCTTTAAGCC	1799
QY	1996	CAAAACGTCCTTAAGCGCTCTCGAGATGAAGACGGGGGAG-CCCCAGCCAGGTGGA	2054
Db	1800	CCAAACGTCCTTAAGCGCTCTCGAGATGAAGACGGGGGAGCCCCAGCCAGGTGGA	1859
QY	2055	GACCCGCGAGAGCGCGCGCGCGGTGTCACGAGGCTCGCACAGCGCGCGCCCTGAG	2114
Db	1860	GACCCGCGAGAGCGCGCGCGCGGTGTCACGAGGCTCGCACAGCGCGCGCCCTGAG	1919
QY	2115	GGTCCGGCC-GAGCCAGGTCCTAAGAGGGCGGCTTGTGTCTCGGGTTAAATAAGGTT	2173
Db	1920	GGTCCGGCCGAGCCAGGTCCTAAGAGGGCGGCTTGTGTCTCGGGTTAAATAAGGTT	1979
QY	2174	CCGTCCG 2180	
Db	1980	CCGTCCG 1986	
RESULT 4			
ABK84431			
ID	ABK84431 standard; cDNA; 2000 BP.		
AC	ABK84431;		
XX			
XX	14-AUG-2002 (first entry)		
DT	Human cDNA differentially expressed in granulocytic cells #1002.		
DE			
XX	Human; ss; granulocytic cell; DNA chip; bacterial infection;		
KW	viral infection; parasitic infection; protozoal infection;		
KW	fungal infection; sterile inflammatory disease; psoriasis;		
KW	rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;		
KW	cardiac reperfusion injury; renal reperfusion injury; ARDS;		
KW	adult respiratory distress syndrome; inflammatory bowel disease;		
KW	Crohn's disease; ulcerative colitis; periodontal disease;		
KW	granulocyte activation; chronic inflammation; allergy.		
OS	Homo sapiens.		
XX	WO200228999-A2.		
FN	11-APR-2002.		
PD			
XX	03-OCT-2001; 2001WO-US030821.		
PF			
XX	03-OCT-2000; 2000US-0237189P.		
PR			
XX	(GENE-) GENE LOGIC INC.		
PA			
XX	Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;		
FI	WPI; 2002-435328/46.		
DR			
XX			

Detecting granulocyte activation by detecting differential expression of genes associated with granulocyte activation, which serves as diagnostic markers that is useful for monitoring disease states and drug toxicity.

Claim 1; SEQ ID NO 1002; 114pp; English.

The invention relates to detecting (M1) granulocyte (GC) activation (GCA) by detecting the level of expression of gene(s) (Gs) identified by DNA chip analysis as given in the specification, and comparing the expression level to an expression level in an unactivated GC, where differential expression of Gs is indicative of GCA. Also included are modulating (M2) GA by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and M5 is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match	89.5%;	Score 1951;	DB 6;	Length 2000;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1984;	Conservative	0;	Mismatches	0; Indels 3; Gaps 3;

QY	196	ATTTTCCTGGGGTCCCGGGCGCGGAGAAAGCTGCATCCAGAGGAGCGCTCCAGGAGCG	255
Db	1	ATTTTCCTGGGGTCCCGGGCGCGGAGAAAGCTGCATCCAGAGGAGCGCTCCAGGAGCG	60
QY	256	GACCCGGAGTGTTCCTCAAGAGCCAGTGAACAAGGACACAGGGGCCCAAGTCCACAGCCAT	315
Db	61	GACCCGGAGTGTTCCTCAAGAGCCAGTGAACAAGGACACAGGGGCCCAAGTCCACAGCCAT	120
QY	316	GCAGACCTGCCCGCTGGCATTCCTTCGCGACGTTTCCAGGCCCTTGGAGCCCTCTGTT	375
Db	121	GCAGACCTGCCCGCTGGCATTCCTTCGCGACGTTTCCAGGCCCTTGGAGCCCTCTGTT	180
QY	376	TTTGCTGCTCTTGTAGTGTCTCAGATGAAGGCTGGGACAGCCCATCTGCACAGAGGG	435
Db	181	TTTGCTGCTCTTGTAGTGTCTCAGATGAAGGCTGGGACAGCCCATCTGCACAGAGGG	240
QY	436	GGTAGTCTCTGTGTCTTGGGGCGAGAACACCGTCTATGTCTTGCAACATCTCCAGCCCTT	495
Db	241	GGTAGTCTCTGTGTCTTGGGGCGAGAACACCGTCTATGTCTTGCAACATCTCCAGCCCTT	300
QY	496	CTCCCATGTCAACATCAAGCTCGTCCCGGAGGAGCGCCATCTTCAATGAGGT	555
Db	301	CTCCCATGTCAACATCAAGCTCGTCCCGGAGGAGCGCCATCTTCAATGAGGT	360
QY	556	GGCTCCAGGCTACTTCTCCCGGGAGCGGTGGCAGCTCCAGGTTTCAGGGAGGCGGTGGCACA	615

Db	1440	GGCGGCGGCTCGGGCTGGGGGGTGAGGTGACGCTGCCTCCGGGGCTGGTTCGGCGCATCCC	1499
Qy	1696	TCAGTCCCTCGGCCACCCCGGGGGTCTCCTCCTCGTGCCACCGCACCTGCGGAGCCTCTTT	1755
Db	1500	TCAGTCCCTCGGCCACCCCGGGGGTCTCCTCCTCGTGCCACCGCACCTGCGGAGCCTCTTT	1559
Qy	1756	TGAGCCAGATCTGTTTCATGCTTTTGTCTTTCTGTCACCTGCGGGGGGGCCCTTTGATGTCCT	1815
Db	1560	TGGACCCAGATCTGTTTCATGCTTTTGTCTTTCTGTCACCTGCGGGGGGGCCCTTTGATGTCCT	1619
Qy	1816	CATCTCTATGCGGTGGAAAAATCACCGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCA	1875
Db	1620	CATCTGTATGCGGTGGAAAAATCACCGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCA	1679
Qy	1876	TGACTCGAATATCTGAAATGAAGAAAAACAAACCGAGCTCACAAACCTCCAAAGTAGCTCCAA	1935
Db	1680	TGACTCGAATATCTGAAATGAAGAAAAACAAACCGAGCTCACAAACCTCCAAAGTAGCTCCAA	1739
Qy	1936	ATGCAATTTTAAAAATGGAATAACAAAAATCTGAAGAAACGTCCTTAAGTGGCTTTTAAGCC	1995
Db	1740	ATGCAATTTTAAAAATGGAATAACAAAAATCTGAAGAAACGTCCTTAAGTGGCTTTTAAGCC	1799
Qy	1996	CCAAACGTCCTTAAGCGCTCTCTCGAGATGAAGACGGGGGGAG-CCCCAGCCAGGTGA	2054
Db	1800	CCAAACGTCCTTAAGCGCTCTCTCGAGATGAAGACGGGGGGAGCCCCCAGCGGTGA	1859
Qy	2055	GACCCCGCAGGACGCGCGCGCGCCGCTGACGAGGCTCGCACAGCGCGGCCCTCTGAG	2114
Db	1860	GACCCCGCAGGACGCGCGCGCGCCGCTGACGAGGCTCGCACAGCGCGGCCCTCTGAG	1919
Qy	2115	GGTCTGGGCC-GAGCCAGGGTCCAAGAGGGGCGCGTTTGTGTCTCGGTTAAAAATAAGTTT	2173
Db	1920	GGTCTGGGCCGAGCCAGGGTCCAAGAGGGGCGCGTTTGTGTCTCGGTTAAAAATAAGTTT	1979
Qy	2174	CCGTCCG 2180	
Db	1980	CCGTCCG 1986	
RESULT 5			
ADJ74845			
ID ADJ74845 standard; DNA; 2000 BP.			
XX	ADJ74845;		
XX	AC		
DT	20-MAY-2004 (first entry)		
XX	Marker gene SEQ ID NO:97.		
DE	bronchial asthma; chronic obstructive pulmonary disease;		
KW	respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;		
KW	gene therapy; marker gene; gene; ds.		
XX	Homo sapiens.		
XX	EP1394274-A2.		
XX	03-MAR-2004.		
XX	04-AUG-2003; 2003EP-00254857.		
XX	06-AUG-2002; 2002JP-00229312.		
PR	20-MAR-2003; 2003JP-00077212.		
XX	(GENO-) GENOX RES INC.		
PA			
XX	Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;		
PI	WPI; 2004-193155/19.		
XX			
XX	Testing for bronchial asthma or chronic obstructive pulmonary disease by		
PT	comparing the expression level of a marker gene in a biological sample		
PT	from a subject with the expression level of the gene in a sample from a		

ADJ74845
ID ADJ74845 standard; DNA; 2000 BP.

XX
DT 20-MAY-2004 (first entry)

XX
KW bronchial asthma; chronic

XX Homo sapiens.
OSXX
PD 03-MAR-2004.

XX 06-AUG-2002; 2002JP-00229312.

PA (GENO-) GENOX RES INC.
XX

DR WPI; 2004-193155/19.
XX

PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT subject with the expression level of the gene in a sample from a
PT subject with the expression level of the gene in a sample from a

PT healthy subject.

XX Claim 1; SEQ ID NO 97; 241pp; English.

XX The present invention describes a method of testing for bronchial asthma or chronic obstructive pulmonary disease. The method comprises determining the expression level of a marker gene in a biological sample from a subject, comparing the expression level determined with the expression level of the marker gene in a biological sample from a healthy subject, and judging whether the subject has bronchial asthma or chronic obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2) whose expression levels decrease when respiratory epithelial cells are stimulated with interleukin-13. Also described: (1) a reagent (I) for testing for bronchial asthma or chronic obstructive pulmonary disease; (2) a kit for screening for a candidate compound for a therapeutic agent to treat bronchial asthma or chronic obstructive pulmonary disease; (3) an animal model for bronchial asthma or chronic obstructive pulmonary disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a method for producing an animal model for bronchial asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease, comprising the compound, a marker gene or an antisense nucleic acid corresponding to a portion of the marker gene, a ribozyme, a polynucleotide that suppresses the expression of the gene through an RNAi effect or an antibody recognising a protein encoded by a marker gene; and (7) a DNA chip for testing for bronchial asthma or a chronic obstructive pulmonary disease, on which a probe has been immobilised to assay a marker gene. (I) has respiratory and antiasthmatic activities, and can be used in gene therapy. The method is useful for testing for or screening for a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease. The present sequence is used in the exemplification of the present invention.

XX Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match		89.58;	Score 1951;	DB 12;	Length 2000;
Best Local Similarity		99.88;	Pred. No. 0;		
Matches 1984;		Conservative 0;	Mismatches 0;	Indels 3;	Gaps 3;
QY	196	ATTTTCTCTGGGCTCCGGGCGGGAGAGCTGCATCCAGAGAGCGGTCAGAGCG	255		
DB	1	ATTTTCTCTGGGCTCCGGGCGGGAGAGCTGCATCCAGAGAGCGGTCAGAGCG	60		
QY	256	GACCCGGAGGTGTTTCAAGACCGAGTGAAGACGAGGAGGAGGAGGAGGAGGAGG	315		
DB	61	GACCCGGAGGTGTTTCAAGACCGAGTGAAGACGAGGAGGAGGAGGAGGAGGAGG	120		
QY	316	GCAGACCTGCCCCCTGGGATTCCTGGCCAGCTTCCAGGCGCTTGGGACCTCTCTGTT	375		
DB	121	GCAGACCTGCCCCCTGGGATTCCTGGCCAGCTTCCAGGCGCTTGGGACCTCTCTGTT	180		
QY	376	TTTGGCTCCCTCTTGAGTGTCTGAGATGAAGGTGGGAGAGGAGGAGGAGGAGGAGG	435		
DB	181	TTTGGCTCCCTCTTGAGTGTCTGAGATGAAGGTGGGAGAGGAGGAGGAGGAGGAGG	240		
QY	436	GGTAGTCTCTGTGCTTGGGCGGAGACCGCTGATGCTCTGCAACATCTCCAAAGCCTT	495		
DB	241	GGTAGTCTCTGTGCTTGGGCGGAGACCGCTGATGCTCTGCAACATCTCCAAAGCCTT	300		
QY	496	CTCCCATGTCAACATCAAGCTGCGTGCCCGGAGGAGGAGGAGGAGGAGGAGGAGG	555		
DB	301	CTCCCATGTCAACATCAAGCTGCGTGCCCGGAGGAGGAGGAGGAGGAGGAGGAGG	360		
QY	556	GGCTCCAGGCTACTTCTCCCGGAGCGGTGGAGCTCAGGTTTGGGAGGCGGTGGCACA	615		
DB	361	GGCTCCAGGCTACTTCTCCCGGAGCGGTGGAGCTCAGGTTTGGGAGGCGGTGGCACA	420		
QY	616	GCTGTGTATCAAGGCGCCCGGAGCTCCCATGCTGGGCTGTACATGTGGCACCCTCGTGG	675		
DB	421	GCTGTGTATCAAGGCGCCCGGAGCTCCCATGCTGGGCTGTACATGTGGCACCCTCGTGG	480		
QY	676	ACACCAGAGAAATAACAGAAAGTCAAGTCAAGTGGAGGTTTCAAGTGCAGAAACCCAGTCCG	735		

DB	481	ACACCAGAGAAATAACAGAAAGTCAAGTCAAGTGGAGGTTTCAAGTGCAGAAACCCAGTCCGC	540		
QY	736	CCCTGACACTGGGTTCTGGCCTGTGCCAGCGGTGTCTACTGTCTTCTATCTCTTGT	795		
DB	541	CCCTGACACTGGGTTCTGGCCTGTGCCAGCGGTGTCTACTGTCTTCTATCTCTTGT	600		
QY	796	CGCTCTGGTCTGTTCCGCTGGTGTACAGTGTGCGCTGTTCCAGCAACACGCGGGAGAGAA	855		
DB	601	CGCTCTGGTCTGTTCCGCTGGTGTACAGTGTGCGCTGTTCCAGCAACACGCGGGAGAGAA	660		
QY	856	GTTCTTCTCTCTAGAAACCCAGATGAAGTGCAGCGCTCTAGAGCGGAGCCAGCAGGG	915		
DB	661	GTTCTTCTCTCTAGAAACCCAGATGAAGTGCAGCGCTCTAGAGCGGAGCCAGCAGGG	720		
QY	916	CTTGAGCAGAGCCTCCGCTGAACTGTGGACCCAGACTCCGAGCCACCCCAAGGCGCT	975		
DB	721	CTTGAGCAGAGCCTCCGCTGAACTGTGGACCCAGACTCCGAGCCACCCCAAGGCGCT	780		
QY	976	GGCACTGGTCTTCAAAACCCCTCACTTTGGAGCCCTGGAGCTGTCTCTCCCCCAACCT	1035		
DB	781	GGCACTGGTCTTCAAAACCCCTCACTTTGGAGCCCTGGAGCTGTCTCTCCCCCAACCT	839		
QY	1036	TGTTTCCATATGCGCAGACCCATAGCCGCTCTCAAGGCGAGAGAGAGAGAGAGAGAG	1095		
DB	840	TGTTTCCATATGCGCAGACCCATAGCCGCTCTCAAGGCGAGAGAGAGAGAGAGAGAG	899		
QY	1096	GCCCTGAGTGGCAGCCTTGGTGGCGGGCTGGTCTCTCTGCTCCACCCGAGGAGCACA	1155		
DB	900	GCCCTGAGTGGCAGCCTTGGTGGCGGGCTGGTCTCTCTGCTCCACCCGAGGAGCACA	959		
QY	1156	GACACCGGCTTGTGTCAGGCTGGGCTCTGTGTACCCACTCTCTGGGTCGTGTCAGAC	1215		
DB	960	GACACCGGCTTGTGTCAGGCTGGGCTCTGTGTACCCACTCTCTGGGTCGTGTCAGAC	1019		
QY	1216	CCTTCCCTTCAACCCCGCAGGCTTCCAGCTCTGCTTCTCAGTTTCCAAATGGAAC	1275		
DB	1020	CCTTCCCTTCAACCCCGCAGGCTTCCAGCTCTGCTTCTCAGTTTCCAAATGGAAC	1079		
QY	1276	ACCTCACCTCCGAGACCCGACTTACAGAGACATGCCCTCTCTCTGCTCTCATCAA	1335		
DB	1080	ACCTCACCTCCGAGACCCGACTTACAGAGACATGCCCTCTCTCTGCTCTCATCAA	1139		
QY	1336	ACCCACAGACCCGAGCTCTCTTTCGCCACCCAGGCTGTGTCGGGCCCGAGGCTGGGGT	1395		
DB	1140	ACCCACAGACCCGAGCTCTCTTTCGCCACCCAGGCTGTGTCGGGCCCGAGGCTGGGGT	1199		
QY	1396	CGGCTCTCTCACTCCAGGCTCCGCGCCCAAGTGAAGGGGGCCCTGCGGAGGCTCAG	1455		
DB	1200	CGGCTCTCTCACTCCAGGCTCCGCGCCCAAGTGAAGGGGGCCCTGCGGAGGCTCAG	1259		
QY	1456	ACACACTGGAGTTTCAAGGCTGGGGGCTTTGGGACATACCTGTCTCTTGGCTATGAGCA	1515		
DB	1260	ACACACTGGAGTTTCAAGGCTGGGGGCTTTGGGACATACCTGTCTCTTGGCTATGAGCA	1319		
QY	1516	GGCTTTGGGGGCTTCCGCGGAGCCCGGGGCGGAGGTAGGCTCTGGGGGCTTAGAG	1575		
DB	1320	GGCTTTGGGGGCTTCCGCGGAGCCCGGGGCGGAGGTAGGCTCTGGGGGCTTAGAG	1379		
QY	1576	GCTGGATGCTCTCTGGGCCCCACCGCAGGGGGCAAGCGAGGCGGGCTGGAGGCGGC	1635		
DB	1380	GCTGGATGCTCTCTGGGCCCCACCGCAGGGGGCAAGCGAGGCGGGCTGGAGGCGGC	1439		
QY	1636	GGCGGGGCTCGGGGTGGGGGTGAGTGGAGCGCTCCGCGGCTGTGCGGATCCC	1695		
DB	1440	GGCGGGGCTCGGGGTGGGGGTGAGTGGAGCGCTCCGCGGCTGTGCGGATCCC	1499		
QY	1696	TCACTCCCTCGGCGACCCCGGGGTGCTCTCTGTCGCCACCGCACCCTGCGAGCTCTT	1755		
DB	1500	TCACTCCCTCGGCGACCCCGGGGTGCTCTCTCTGTCGCCACCGCACCCTGCGAGCTCTT	1559		
QY	1756	TGGAACCAATCTGTTTCTGTTCTGCTCTGCTGGGGGGCTTGTGTTGTTGTTGTT	1815		

Db 1560, TGGACCCAGATCTGTTTCATGCTTTTGTCTTCCTCACTGCGGGGGGCGCTTTGATGTCCT 1619
Qy 1816 CATCTGTATGGGTGAAAAATCACCGGAATCCCGCTTTCAGTCTTTTGAAAAAGTTCCA 1875
Db 1620 CATCTGTATGGGTGAAAAATCACCGGAATCCCGCTTTCAGTCTTTTGAAAAAGTTCCA 1679
Qy 1876 TGACTCGAATATCTGAATGAAGAAAAAACAACCGACTCAAACTCTCAAGTAGTCTCAA 1935
Db 1680 TGACTCGAATATCTGAATGAAGAAAAAACAACCGACTCAAACTCTCAAGTAGTCTCAA 1739
Qy 1936 ATGCAATTTTAAATGGAACAAAAATCTGAAGAAACGCTTTAGTGGCTTTAAGCC 1995
Db 1740 ATGCAATTTTAAATGGAACAAAAATCTGAAGAAACGCTTTAGTGGCTTTAAGCC 1799
Qy 1996 CCAAAACGTCCTTAAGGCGTCTCGAGATGAAGAGGGGGGAG-CCCCAGCAGGTGGA 2054
Db 1800 CCAAAACGTCCTTAAGGCGTCTCGAGATGAAGAGGGGGGAGCCCCAGCAGGTGGA 1859
Qy 2055 GACCCGCGAGAGCGGGCGGCCCGGTGACCGAGGCTTCGCACAGCGCGCGCCCTGAG 2114
Db 1860 GACCCGCGAGAGCGGGCGGCCCGGTGACCGAGGCTTCGCACAGCGCGCGCCCTGAG 1919
Qy 2115 GGTCCGGCC-GAGCCAGGCTCCAAGAGGGCGGCTTTGTCTCGGTTAAATTAAGCTT 2173
Db 1920 GGTCCGGCCGAGCCAGGCTCCAAGAGGGCGGCTTTGTCTCGGTTAAATTAAGCTT 1979
Qy 2174 CCGTCCG 2180
Db 1980 CCGTCCG 1986

RESULT 6
ADR52746
ID ADR52746 standard; DNA; 2000 BP.
XX ADR52746;
AC ADR52746;
DT 18-NOV-2004 (first entry)
XX
XX Drug therapy altered expressed gene #97.
XX
KW drug activity monitoring; expression profile; gene expression;
KW peripheral blood sample; peripheral blood mononuclear cell; drug therapy;
KW CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;
KW mTOR; ds.
XX
OS Homo sapiens.
PN WO2004072265-A2.
XX
PD 26-AUG-2004.
XX
PF 11-FEB-2004; 2004WO-US004118.
XX
PR 11-FEB-2003; 2003US-0446133P.
PR 03-APR-2003; 2003US-0459782P.
PR 23-JAN-2004; 2004US-0538246P.
XX
XX (AMHP) WYETH.
PA (BURC/) BURCZYNSKI M.
PA (TWIN/) TWINE N.
PA (DORN/) DORNER A. J.
PA (TREP/) TREPICCHIO W. L.
XX
PI Burczynski M, Twine N, Dorner AJ, Trepicchio WL;
XX WPI; 2004-642301/62.
XX
XX Monitoring drug activities in vivo comprises comparing an expression
PT profile of a gene in a peripheral blood sample of a patient before and
PT after drug therapy.
XX
PS Disclosure; SEQ ID NO 97; 136bp; English.

XX The invention relates to a method of monitoring drug activities in vivo
CC by comparing an expression profile of at least one gene in a peripheral
CC blood sample of a patient to a reference expression profile of the at
CC least one gene, where the at least one gene is differentially expressed
CC in peripheral blood mononuclear cells (PBMCs) of patients who have a non-
CC blood disease and are subjected to a drug therapy as compared to PBMCs
CC isolated from the patient before the drug therapy, and where the patient
CC has the non-blood disease and is being treated by the drug therapy. The
CC method, kit, and nucleic acid array are useful for monitoring drug
CC activities in vivo. The drug is especially CCI-779, an ester analogue of
CC the immunosuppressant rapamycin which is a potent inhibitor of the
CC mammalian target of rapamycin (mTOR). This sequence represents a gene
CC expressed in PBMC altered by the drug therapy. (Note: this sequence does
CC no form part of the printed specification but was obtained in electronic
CC format from WIPO at ftp.wipo.int/pub/published_pct_sequences/).

XX Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;
SQ

Query Match 89.5%; Score 1951; DB 13; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

Qy 196 ATTTTCTTGGGCTCCGGGGCGGAGAAAGCTGATCCAGAGAGCGGTCCAGGAGCG 255
Db 1 ATTTTCTTGGGCTCCGGGGCGGAGAAAGCTGATCCAGAGAGCGGTCCAGGAGCG 60
Qy 256 GACCCGGGAGTGTTCAGAGCCAGTGACAGAGGACCCAGGGGCCCAAGTCCACAGCCAT 315
Db 61 GACCCGGGAGTGTTCAGAGCCAGTGACAGAGGACCCAGGGGCCCAAGTCCACAGCCAT 120
Qy 316 GCAGACCTGCCCTTGGGCTCCGTCATTCCTGCGCCAGCTTTCCAGGCGCTTGGGAGCCCTCTGT 375
Db 121 GCAGACCTGCCCTTGGGCTCCGTCATTCCTGCGCCAGCTTTCCAGGCGCTTGGGAGCCCTCTGT 180
Qy 376 TTTGGCTGCTTCTTGGGCTCCAGATGAAGGCTGGGAGACGCCCATCTTCACAGAGGG 435
Db 181 TTTGGCTGCTTCTTGGGCTCCAGATGAAGGCTGGGAGACGCCCATCTTCACAGAGGG 240
Qy 436 GGTAGTCTCTGCTTGGGCGGAGACACCGTCTCATGTCCTGCAACATCTCCACGCTT 495
Db 241 GGTAGTCTCTGCTTGGGCGGAGAACACCGTCTCATGTCCTGCAACATCTCCACGCTT 300
Qy 496 CTCCTCATGTCAACATCAAGCTGCGTCCCGGAGGAGAGCGGCATCTTCAATGAGGT 555
Db 301 CTCCTCATGTCAACATCAAGCTGCGTCCCGGAGGAGAGCGGCATCTTCAATGAGGT 360
Qy 556 GGCTCCAGGCTACTTCTCCCGGAGCGGTGCGAGCTCCAGGTTTCAAGGAGCGGTGGCACA 615
Db 361 GGCTCCAGGCTACTTCTCCCGGAGCGGTGCGAGCTCCAGGTTTCAAGGAGCGGTGGCACA 420
Qy 616 GCTGTGATCAAAAGCGCGCGGACTCCCATGCTGGGCTGTACATGTGGCACCCTCGTGGG 675
Db 421 GCTGTGATCAAAAGCGCGCGGACTCCCATGCTGGGCTGTACATGTGGCACCCTCGTGGG 480
Qy 676 ACACGAGAGAAATAACAGACAAGTCAAGCTCGAGGTTTTCAGGTGCAGAACCCAGTCCGC 735
Db 481 ACACGAGAGAAATAACAGACAAGTCAAGCTCGAGGTTTTCAGGTGCAGAACCCAGTCCGC 540
Qy 736 CCCTGACACTGGGTTCTGGGCTGTGCGAGCGGTGTCACTGCTGTTCATCTCTTGGT 795
Db 541 CCCTGACACTGGGTTCTGGGCTGTGCGAGCGGTGTCACTGCTGTCTTCTCTTGGT 600
Qy 796 GGCTCTGTGATCTCCGCTGTTACAGTGGTGGCTGTTCAGGAGCGCGGGGAGAGAA 855
Db 601 GGCTCTGTGATCTCCGCTGTTACAGTGGTGGCTGTTCAGGAGCGCGGGGAGAGAA 660
Qy 856 GTTCTTCTCTAGAACCCAGAGTGAAGGTCCAGGCGCTTCAGAGCGGAGCGCCAGCAGGG 915
Db 661 GTTCTTCTCTAGAACCCAGAGTGAAGGTCCAGGCGCTTCAGAGCGGAGCGCCAGCAGGG 720
Qy 916 CCTGAGCAGAGCTCCGCTGAACTGTGACCCCCAGACTCCAGGCGCCACCCCAAGGCCGCT 975

Db	721	CCTGAGCAGAGCCTCCGCTGAACCTGTGACCCCGACACTCCGAGCCACCCCAAGCCGCT	780
Qy	976	GGCACTGGTGTTCACAAACCCCTCACCACCTTGGAGCCCTGGAGCTGTGTCTCCCCCAACCCCT	1035
Db	781	GGCACTGGTGTTCACAAACCCCTCACCACCTTGGAGCCCTGGAGCTGTGT-CCCCCAACCCCT	839
Qy	1036	TGTTTCCATATGCGGAGACCCATAGCCGCTGCAAGGAGAGGACACAGGAGGCCA	1095
Db	840	TGTTTCCATATGCGGAGACCCATAGCCGCTGCAAGGAGAGGACACAGGAGGCCA	899
Qy	1096	GGCCTGAGTCCGACCTTGGGTGGCGGCGCTGGGTCTCTGCTCCACCCCGAGAGGCACA	1155
Db	900	GGCCTGAGTCCGACCTTGGGTGGCGGCGCTGGGTCTCTGCTCCACCCCGAGAGGCACA	959
Qy	1156	GACACGGCTTCTTGGCAGGCTGGGCTCTGTGTACCCACTCTCTGGGTGGTGCAGAC	1215
Db	960	GACACGGCTTCTTGGCAGGCTGGGCTCTGTGTACCCACTCTCTGGGTGGTGCAGAC	1019
Qy	1216	CCTTCCCTCCACCCCGCAGCTTCTCCAGCTCTGCTTCCAGTCTCTGCTCCAGTCTTCCAAATGGAACC	1275
Db	1020	CCTTCCCTCCACCCCGCAGCTTCTCCAGCTCTGCTTCCAGTCTCTGCTCCAGTCTTCCAAATGGAACC	1079
Qy	1276	ACCTCAGCTCCGAGCACCAGCTTACAGAGCAGATGCCCTCCCTCTGCTGCTCATCAA	1335
Db	1080	ACCTCAGCTCCGAGCACCAGCTTACAGAGCAGATGCCCTCCCTCTGCTGCTCATCAA	1139
Qy	1336	ACCACAGACCCGAGCTCCCTTCTGCAACCCAGGCTGCTCGGCGCCAGGTGTGGGT	1395
Db	1140	ACCACAGACCCGAGCTCCCTTCTGCAACCCAGGCTGCTCGGCGCCAGGTGTGGGT	1199
Qy	1396	CCGCTCTCTCCACTCCAGGCTCCGCGCCAGTGGAGGGGCCCTGCGGAGGCTCAG	1455
Db	1200	CCGCTCTCTCCACTCCAGGCTCCGCGCCAGTGGAGGGGCCCTGCGGAGGCTCAG	1259
Qy	1456	ACACACTGGAGTTTCCAGGCTGGGGGGCTTGGCACAATACCTGTCTGCTTGGTATGACA	1515
Db	1260	ACACACTGGAGTTTCCAGGCTGGGGGGCTTGGCACAATACCTGTCTGCTTGGTATGACA	1319
Qy	1516	GGCTTTGGGGGGCTTCCGGGGCAGCCCGGGGGCCGAGGTAGGTCTGGGGCTTAGAG	1575
Db	1320	GGCTTTGGGGGGCTTCCGGGGCAGCCCGGGGGCCGAGGTAGGTCTGGGGCTTAGAG	1379
Qy	1576	GCTGGGATGCTCTCTGCCCCCAGCGCCAGGGGCAAGCGAGGCGGGCTGGAGGCGGC	1635
Db	1380	GCTGGGATGCTCTCTGCCCCCAGCGCCAGGGGCAAGCGAGGCGGGCTGGAGGCGGC	1439
Qy	1636	GGCGGGGCTCGGGTGGGGGTGAGGTGACGCTGCTCCGGGGTGTGCGGATCCC	1695
Db	1440	GGCGGGGCTCGGGTGGGGGTGAGGTGACGCTGCTCCGGGGTGTGCGGATCCC	1499
Qy	1696	TCAGTCCCTCGGCACCCGGGGTTCGCTCCCTCGTCCACCGCAGCTGCGGAGCCTCTT	1755
Db	1500	TCAGTCCCTCGGCACCCGGGGTTCGCTCCCTCGTCCACCGCAGCTGCGGAGCCTCTT	1559
Qy	1756	TGGACCCAGATCTCTTCATGCTTTGCTTCGTCACTGCGCGGGGCCCTTTGATGCTT	1815
Db	1560	TGGACCCAGATCTCTTCATGCTTTGCTTCGTCACTGCGCGGGGCCCTTTGATGCTT	1619
Qy	1816	CATCTGTATGGGTGGAAAAATACCGGGGAATCCCTTCAGTCTTTGAAAAAGTTCCA	1875
Db	1620	CATCTGTATGGGTGGAAAAATACCGGGGAATCCCTTCAGTCTTTGAAAAAGTTCCA	1679
Qy	1876	TGACTCGAATCTGAATGAAGAAACAAACCGACTCACAACTCCCAAGTAGTCCAA	1935
Db	1680	TGACTCGAATCTGAATGAAGAAACAAACCGACTCACAACTCCCAAGTAGTCCAA	1739
Qy	1936	ATGCAATTTTAAATGAAACAAAAATCTGAAAGAAACGCTTTTGTGGCTTTAAGCC	1995
Db	1740	ATGCAATTTTAAATGAAACAAAAATCTGAAAGAAACGCTTTTGTGGCTTTAAGCC	1799
Qy	1996	CCAAACGCTCCCTAAGGCTCTCGAGTGAAGACGGGGGGAG-CCCCAGCCAGGTGGA	2054
Db	1800	CCAAACGCTCCCTAAGGCTCTCGAGTGAAGACGGGGGGAGCCCCCAGCCAGGTGGA	1859

Qy	2055	GACCCCGCAGGACGCGCGCGCCCGGTGACCCGAGGCTTCGACACGCCGCGCCCTGAG	2114
Db	1860	GACCCCGCAGGACGCGCGCGCCCGGTGACCCGAGGCTTCGACACGCCGCGCCCTGAG	1919
Qy	2115	GGTCGGGCC- GAGCCAGGGTCCAAAGGGGCGGCTTTGTCTCGGGTTAAATAAGGTT	2173
Db	1920	GGTCGGGCCGAGCCAGGGTCCAAAGGGGCGGCTTTGTCTCGGGTTAAATAAGGTT	1979
Qy	2174	CCGTCCG 2180	
Db	1980	CCGTCCG 1986	

RESULT 7
ADP24663
ID ADP24663 standard; cDNA; 2000 BP.
XX
AC ADP24663;
XX
DT 18-NOV-2004 (first entry)
XX
DE PRO polypeptide encoding cDNA SEQ ID NO:1841.
XX
ss; gene; PRO; antinflammatory; antiarthritic; antirheumatic;
KW immunosuppressive; osteopathic; antidiabetic; dermatological;
KW antipruritic; antiallergic; antiasthmatic; hepatotropic; respiratory;
KW gene therapy; immune system.
XX
OS Unidentified.
XX
PN WO2004041170-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034312.
XX
PR 01-NOV-2002; 2002US-0423394P.
XX
PA (GETH) GENENTECH INC.
XX
PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WT;
PI Wu TD;
XX
WFI; 2004-419628/39.
P-PSDB; ADP24664.
XX
New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
Claim 1; SEQ ID NO 1841; 2940pp; English.
XX
The invention relates to a novel isolated nucleic acid and the PRO
polypeptide encoded by it. A protein of the invention has
CC antinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipruritic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, autoimmune
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's

CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence encodes a PRO protein of
CC the invention.
XX

SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 89.5%; Score 1951; DB 13; Length 2000;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY	196	ATTTCCTCGGGCTCCGGGCGGAGAGAGCTGCATCCAGAGGAGCGGTCCAGAGCG	255
DB	1	ATTTCCTCGGGCTCCGGGCGGAGAGAGCTGCATCCAGAGGAGCGGTCCAGAGCG	60
QY	256	GACCCGGAGTGTTCAGAGCAGTGAACAAGACAGAGGCGCCAAAGTCCACAGCCAT	315
DB	61	GACCCGGAGTGTTCAGAGCAGTGAACAAGACAGAGGCGCCAAAGTCCACAGCCAT	120
QY	316	GCAGACCTGCCCTCGGATTCCTCGGCCAGCTTTCCAGGCGCTTGGAGCCCTCTGT	375
DB	121	GCAGACCTGCCCTCGGATTCCTCGGCCAGCTTTCCAGGCGCTTGGAGCCCTCTGT	180
QY	376	TTTGGCTGCTCTTGAAGTCTCAGAAATGAAGGCTGGAGACGCCCATCTGCACAGAGG	435
DB	181	TTTGGCTGCTCTTGAAGTCTCAGAAATGAAGGCTGGAGACGCCCATCTGCACAGAGG	240
QY	436	GATAGTCTCTGTGCTTGGGGCGAGAACACCGTATGCTCTGCAACATCTCCAAAGCCTT	495
DB	241	GATAGTCTCTGTGCTTGGGGCGAGAACACCGTATGCTCTGCAACATCTCCAAAGCCTT	300
QY	496	CTCCATGTCACATCAAGTCTGCTGCGCCACCGGAGGAGCGCATCTTCAATGAGT	555
DB	301	CTCCATGTCACATCAAGTCTGCTGCGCCACCGGAGGAGCGCATCTTCAATGAGT	360
QY	556	GGCTCCAGGCTACTTCTCCGGGAGCGGCTGGCAGCTCCAGGTTCAAGGAGGCGTGCCACA	615
DB	361	GGCTCCAGGCTACTTCTCCGGGAGCGGCTGGCAGCTCCAGGTTCAAGGAGGCGTGCCACA	420
QY	616	GCTGTGATCAAAAGCGCCCGGACTCCCATGCTGGGCTGTACATGTGCACTCTGTGG	675
DB	421	GCTGTGATCAAAAGCGCCCGGACTCCCATGCTGGGCTGTACATGTGCACTCTGTGG	480
QY	676	ACACAGAGAAATACAGACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	735
DB	481	ACACAGAGAAATACAGACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	540
QY	736	CCCTGACACTGGGTTCTGGCTGTGTCAGCGGTGGTCACTGTCTTTCATCTCTTGT	795
DB	541	CCCTGACACTGGGTTCTGGCTGTGTCAGCGGTGGTCACTGTCTTTCATCTCTTGT	600
QY	796	CGCTCTGTGTCATGTCGCTGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	855
DB	601	CGCTCTGTGTCATGTCGCTGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	660
QY	856	GTTCCTCTCTAGAACCCAGATGAGTTCAGCGCTTCAGCGGCGGAGCCACAGCG	915
DB	661	GTTCCTCTCTAGAACCCAGATGAGTTCAGCGCTTCAGCGGCGGAGCCACAGCG	720
QY	916	CTTGAGCAGAGCTCTCGCTGAAGTGTGGACCCAGACTCCAGGCGCCACCCCAAGCCGCT	975
DB	721	CTTGAGCAGAGCTCTCGCTGAAGTGTGGACCCAGACTCCAGGCGCCACCCCAAGCCGCT	780
QY	976	GGCACTGGTGTCAAAACCTCAACCTTGGAGCCCTGGAGTGTCTCTCCCGCCCAACCT	1035
DB	781	GGCACTGGTGTCAAAACCTCAACCTTGGAGCCCTGGAGTGTCTCTCCCGCCCAACCT	839
QY	1036	TGTTTCCATATGCCGAGACCCATAGCCGCTTGAAGGAGAGAGGACACAGGAGGCA	1095

DB	840	TGTTTCCATATGCCGAGACCCATAGCCGCTTGAAGGAGAGAGGACACAGGAGGCA	899
QY	1096	GGCTGAGTGGCGGCTCTTGGGTGGCGGGGCTGGGTCTCTCGCTCCACCCGAGGAGGCA	1155
DB	900	GGCTGAGTGGCGGCTCTTGGGTGGCGGGGCTGGGTCTCTCGCTCCACCCGAGGAGGCA	959
QY	1156	GACACCGGCTTGTCTTGGCAGGCTGGGCTCTGTGTACCCACTCTCTGGGTGGGTGAGAC	1215
DB	960	GACACCGGCTTGTCTTGGCAGGCTGGGCTCTGTGTACCCACTCTCTGGGTGGGTGAGAC	1019
QY	1216	CGTTCCTCCACCCCGGCTCTTCCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1275
DB	1020	CGTTCCTCCACCCCGGCTCTTCCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1079
QY	1276	ACCTCACCTCCGAGACCCGACTTACAGGACGATGCCCTCTCTCTCTCTCTCTCTCTCT	1335
DB	1080	ACCTCACCTCCGAGACCCGACTTACAGGACGATGCCCTCTCTCTCTCTCTCTCTCTCT	1139
QY	1336	ACCCAGACCCGAGCTCTCTTCTGCGACCCAGGCTGGTCCGCGCCAGGTGTGGGT	1395
DB	1140	ACCCAGACCCGAGCTCTCTTCTGCGACCCAGGCTGGTCCGCGCCAGGTGTGGGT	1199
QY	1396	CGGCTCTCTCACCTCCAGGCTCCGCGCCCAAGTGAGGGGCGCTCTGCGAGGCTCAG	1455
DB	1200	CGGCTCTCTCACCTCCAGGCTCCGCGCCCAAGTGAGGGGCGCTCTGCGAGGCTCAG	1259
QY	1456	ACACACTGGAGTTCAGGGCTGGGGGGCTTGGGCACATACCTGTCTCTCTCTCTCTCTCT	1515
DB	1260	ACACACTGGAGTTCAGGGCTGGGGGGCTTGGGCACATACCTGTCTCTCTCTCTCTCTCT	1319
QY	1516	GGCTTGGGGGCTTCTCCGCGGAGCCCGGGGGCGAGGTAGGTCTTGGGGCTTAGAG	1575
DB	1320	GGCTTGGGGGCTTCTCCGCGGAGCCCGGGGGCGAGGTAGGTCTTGGGGCTTAGAG	1379
QY	1576	GCTGGGATGGCTCTTGGCGCCACCGCAGGGGGGCAAGCGCGGGCTGGGAGGCGCG	1635
DB	1380	GCTGGGATGGCTCTTGGCGCCACCGCAGGGGGGCAAGCGCGGGCTGGGAGGCGCG	1439
QY	1636	GGCGGGCTCGGGCTGGGGGGCTCAGGTGGAGCTGCTCTCGGGGCTGGTCCCGCATCC	1695
DB	1440	GGCGGGCTCGGGCTGGGGGGCTCAGGTGGAGCTGCTCTCGGGGCTGGTCCCGCATCC	1499
QY	1696	TCAGTCCCTCGGCGACCCCGGGGGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1755
DB	1500	TCAGTCCCTCGGCGACCCCGGGGGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1559
QY	1756	TGGACCCAGATCTGTTCATGCTTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1815
DB	1560	TGGACCCAGATCTGTTCATGCTTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1619
QY	1816	CATCTGTATGGGTGGAAAAATCACCGGGAATCCCTCTCTCTCTCTCTCTCTCTCTCTCT	1875
DB	1620	CATCTGTATGGGTGGAAAAATCACCGGGAATCCCTCTCTCTCTCTCTCTCTCTCTCTCT	1679
QY	1876	TGACTCGAATATCTGAAATGAAGAAAAAACAACCGACTCAAAAACCTCCAAAGTAGTCCAA	1935
DB	1680	TGACTCGAATATCTGAAATGAAGAAAAAACAACCGACTCAAAAACCTCCAAAGTAGTCCAA	1739
QY	1936	ATGCAATTTTAAATGGAACAAAAATCTGAAGAAACGCTCTTTAGTGGCTTTAAGCC	1995
DB	1740	ATGCAATTTTAAATGGAACAAAAATCTGAAGAAACGCTCTTTAGTGGCTTTAAGCC	1799
QY	1996	CCAAAACT	2054
DB	1800	CCAAAACT	1859
QY	2055	GACCCCGAGAGCGCGCGCGCGGTGAACGAGGCTTCGACAGCGCGCGCGCGCTGAG	2114
DB	1860	GACCCCGAGAGCGCGCGCGCGGTGAACGAGGCTTCGACAGCGCGCGCGCGCTGAG	1919
QY	2115	GCTGGGGC-CAGCCAGGCTCCAGAGGGGCGCTTGTGTCTCTGGGTAAAAATAGGTT	2173
DB	1920	GCTGGGGC-CAGCCAGGCTCCAGAGGGGCGCTTGTGTCTCTGGGTAAAAATAGGTT	1979

CC organic molecule to a cell that expresses the protein described above.
 CC The TAT sequences have cytostatic activities, and can be used in gene
 CC therapy. The composition and methods are useful for diagnosing,
 CC preventing or treating cancer. The composition is also used for preparing
 CC a medicament for the therapeutic treatment or diagnostic detection of a
 CC cell proliferative disorder or cancer. The present sequence represents a
 CC human TAT cDNA sequence from the present invention.
 SQ

XX Sequence 1985 BP; 372 A; 649 C; 594 G; 370 T; 0 U; 0 Other;

Query Match 80.3%; Score 1751; DB 12; Length 1985;
 Best Local Similarity 97.8%; Pred. No. 0;
 Matches 1944; Conservative 0; Mismatches 25; Indels 18; Gaps 16;

QY 196 ATTTTCTGGGGCTCCGGGCGGAGAGTCTCATCCAGAGAGCGGTCCAGAGCG 255
 Db 1 ATTTTCTGGGGCTCCGGGCGGAGAGTCTCATCCAGAGAGCGGTCCAGAGCG 60
 QY 256 GACCCGGGAGTGTTCAGAGAGCCAGTGAAGAACACGAGGCGCCAAAGTCCACAGCCAT 315
 Db 61 GACCCGGGAGTGTTCAGAGAGCCAGTGAAGAACACGAGGCGCCAAAGTCCACAGCCAT 119
 QY 316 GCAGACCTGCCCCCTGGCATTCTTGGCCACAGTTCCTCCAGGCGCTTGGGACCCCTCTGT 375
 Db 120 GCAGACCTGCCCCCTGGCATTCTTGGCCACAGTTCCTCCAGGCGCTTGGGACCCCTCTGT 179
 QY 376 TTTGGCTGCTCTTGAAGTGTCTAGATGAAGGCTGGGACAGCCCATCTTGCACAGAGG 435
 Db 180 TTTGGCTGCTCTTGAAGTGTCTAGATGAAGGCTGGGACAGCCCATCTTGCACAGAGG 239
 QY 436 GGTAGTCTCTGTGTTGGGGGAGAAACACGCTCATGCTCTGCACATCTCCACGCGCTT 495
 Db 240 GGTAGTCTCTGTGTTGGGGGAGAAACACGCTCATGCTCTGCACATCTCCACGCGCTT 299
 QY 496 CTCCCATGTCAACATCAAGCTGCGTCCAGGCGAGGAGCGCCATCTTCAATGAGGT 555
 Db 300 CTCCCATGTCAACATCAAGCTGCGTCCAGGCGAGGAGCGCCATCTTCAATGAGGT 359
 QY 556 GGCTCCAGGCTACTTCTCCGGGACGCTGGCAGCTCCAGGTTGAGGAGCGTGGCACA 615
 Db 360 GGCTCCAGGCTACTTCTCCGGGACGCTGGCAGCTCCAGGTTGAGGAGCGTGGCACA 419
 QY 616 GCTGTGATCAAGGCGCGCGGCTCCATGCTGGGCTGTACATGTGGACCTCTGTGG 675
 Db 420 GCTGTGATCAAGGCGCGCGGCTCCATGCTGGGCTGTACATGTGGACCTCTGTGG 479
 QY 676 ACACAGAGAGAAATACAGACAAAGTCAAGTGGAGGTTTTCAGTGCAGAAACCCAGTCCGC 735
 Db 480 ACACAGAGAGAAATACAGACAAAGTCAAGTGGAGGTTTTCAGTGCAGAAACCCAGTCCGC 539
 QY 736 CCTGACACTGGGTTCTGGCTGTGCGAGCGGTGTCTGCTGCTTCTTCTTCTTCTT 795
 Db 540 CCTGACACTGGGTTCTGGCTGTGCGAGCGGTGTCTGCTGCTTCTTCTTCTTCTT 598
 QY 796 CGCTCTGTCATGTTGCGCTGTGTCAGAGTGGCGCTGTTCCTCCAGCAACCGCGAGAGAA 855
 Db 599 CGCTCTGTCATGTTGCGCTGTGTCAGAGTGGCGCTGTTCCTCCAGCAACCGCGAGAGAA 657
 QY 856 GTTCTTCTCTCTAGAAACCCCAAGTGAAGTTCAGGCGCTTCAGAGCGGAGCGCCAGCGG 915
 Db 658 GTTCTTCTCTCTAGAAACCCCAAGTGAAGTTCAGGCGCTTCAGAGCGGAGCGCCAGCGG 716
 QY 916 CTGAGCAGAGCTCGCTGTAAGTGGACCCCAAGTTCAGGCGCGCCAGCGCGCGCT 975
 Db 717 CTGAGCAGAGCTCGCTGTAAGTGGACCCCAAGTTCAGGCGCGCGCT 775
 QY 976 GGCACCTGGTGTTCAAACCCCTCAACCTTGGAGCGCTTGGAGCTGTGTCTCCCGCCACCT 1035
 Db 776 GGCACCTGGTGTTCAAACCCCTCAACCTTGGAGCGCTTGGAGCTGTGTCTCCCGCCACCT 834
 QY 1036 TGTTCATATGCGGAGAGCCCATAGCCGCTTGCAGGAGAGAGAGAGAGAGAGAGAGAG 1095
 Db 835 TGTTCATATGCGGAGAGCCCATAGCCGCTTGCAGGAGAGAGAGAGAGAGAGAGAGAG 894

RESULT 8
 ADQ86642
 ID ADQ86642 standard; cDNA; 1985 BP.
 XX
 AC ADQ86642;
 XX
 DT 07-OCT-2004 (first entry)
 XX
 DE Human tumour-associated antigenic target (TAT) cDNA sequence #3516.
 XX
 KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
 KW cancer; cell proliferative disorder; gene; ss.
 OS Homo sapiens.
 XX
 PN WO2004060270-A2.
 XX
 PD 22-JUL-2004.
 XX
 PF 15-OCT-2003; 2003WO-US029126.
 XX
 PR 18-OCT-2002; 2002US-0418988P.
 XX
 XX (GETH) GENENTECH INC.
 PA (WUTD/) WU T D.
 PA (ZHOU/) ZHOU Y.
 XX
 XX Wu TD, Zhou Y;
 XX
 XX WPT; 2004-534300/51.
 DR
 PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
 PT preventing or treating cell proliferative disorders such as cancer.
 XX
 PS Claim 1; SEQ ID NO 3516; 5504pp; English.

CC The present invention describes an isolated tumour-associated antigenic
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
 CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
 CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
 CC (c). Also described: (1) an expression vector comprising the above
 CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
 CC a process for producing a polypeptide; (4) an isolated polypeptide
 CC comprising: (a) an amino acid sequence encoded by any of the above
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
 CC length coding region of the above nucleotide sequences; or (c) a sequence
 CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
 CC an isolated antibody that binds to the above polypeptide; (7) a process
 CC for producing the antibody; (8) an isolated oligopeptide that binds to
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
 CC binding organic molecule that binds to the above polypeptide; (10) a
 CC composition of matter comprising the above (chimeric) polypeptide,
 CC antibody, oligopeptide or TAT binding organic molecule, in combination
 CC with a carrier; (11) an article of manufacture comprising a container and
 CC the composition of matter contained within the container; (12) methods of
 CC inhibiting the growth of a cell that expresses the above protein, where
 CC the growth of the cell is at least in part dependent upon a growth
 CC potentiating effect of the above protein; (13) a method of
 CC therapeutically treating a mammal having a cancerous tumour comprising
 CC cells that express the above protein; (14) a method of determining the
 CC presence of a protein in a sample suspected of containing the protein
 CC described above; (15) methods of diagnosing the presence of a tumour in a
 CC mammal; (16) a method for treating or preventing a cell proliferative
 CC disorder associated with increased expression or activity of the above
 CC protein; and (17) a method of binding an antibody, oligopeptide or

CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 1985 BP; 372 A; 649 C; 594 G; 370 T; 0 U; 0 Other;

Query Match 80.3%; Score 1751; DB 13; Length 1985;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 1944; Conservative 0; Mismatches 25; Indels 18; Gaps 16;

QY 196 ATTTCCTCGGGCTCCGGGCGCGGAGAGCTGCATCCAGAGGAGCGCTCCAGGAGCG 255
DB 1 ATTTCCTCGGGCTCCGGGCGCGGAGAGCTGCATCCAGAGGAGCGCTCCAGGAGCG 60

QY 256 GACCCGGGAGTGTTCAGAGCCAGTGACAGGAGCCAGGGGCCCAAGTCCACACAGCCAT 315
DB 61 GACCCGGGAGTGTTCAGAGCCAGTGACAGGAGCCAGGGGCCCAAGTCCACACAGCCAT 119

QY 316 GCAGACTCGCCCTCGGCATTCCTGGCCACGCTTCCAGGCGCTTGGAGACCTCTCTGTT 375
DB 120 GCAGACTCGCCCTCGGCATTCCTGGCCACGCTTCCAGGCGCTTGGAGACCTCTCTGTT 179

QY 376 TTTGGTGTCTCTGAGTGTCTAGAAATGAAGCTGGGACAGCCCATCTGCACAGAGGG 435
DB 180 TTTGGTGTCTCTGAGTGTCTAGAAATGAAGCTGGGACAGCCCATCTGCACAGAGAGG 239

QY 436 GGTAGTCTGTCTTGGGGGAGAACACCGTCATGTCTCTGCAACATCTCCAAAGCCCTT 495
DB 240 GGTAGTCTGTCTTGGGGGAGAACACCGTCATGTCTCTGCAACATCTCCAAAGCCCTT 299

QY 496 CTCCTCATGTCAACATCAAGTGTGTCGCCAGGAGAGAGGCCCATCTTCAATGAGGT 555
DB 300 CTCCTCATGTCAACATCAAGTGTGTCGCCAGGAGAGAGGCCCATCTTCAATGAGGT 359

QY 556 GGCTCCAGGCTACTTCTCCGGAGCGCTGGAGCTCCAGGTTTCAGGGAGGGGTGGCACA 615
DB 360 GGCTCCAGGCTACTTCTCCGGAGCGCTGGAGCTCCAGGTTTCAGGGAGGGGTGGCACA 419

QY 616 GCTGGTGTATCAAAAGCGCCCGGACTCCCATGTGGGCTGTATCATGTGGACCTCTGGG 675
DB 420 GCTGGTGTATCAAAAGCGCCCGGACTCCCATGTGGGCTGTATCATGTGGACCTCTGGG 479

QY 676 ACACAGAGAAATACAGACAGTCAAGTCAAGTGTGGAGTTTCAGGTCGAGAACCCAGTCCGC 735
DB 480 ACACAGAGAAATACAGACAGTCAAGTCAAGTGTGGAGTTTCAGGTCGAGAACCCAGTCCGC 539

QY 736 CCCTGACACTGGGTCTGGGCTGTGCCAGGGTGGTCACTGCTGTCTTTCATCTCTTGGT 795
DB 540 CCCTGACACTGGGTCTGGGCTGTGCCAGGGTGGTCACTGCTGTCTTTCATCTCTTGGT 598

QY 796 CGCTCTGTGTATGTTCGGCTCGGTACAGTGGCCGTGTTCAGCAACCGCGGAGAGAA 855
DB 599 CGCTCTGTGTATGTTCGGCTCGGTACAGTGGCCGTGTTCAGCAACCGCGGAGAGAA 657

QY 856 GTTCTTCTCTAGAACCCAGATGAGGTGCGAGCCCTCAGAGCGGAGCCACAGCAGGG 915
DB 658 GTTCTTCTCTAGAACCCAGATGAGGTGCGAGCCCTCAGAGCGGGA-CGACAGCAGGG 716

QY 916 CCTGAGCAGAGCCTCCGCTGAACCTGTGGACCCCGAGCTCCAGGCCACCCCAAGCGCGCT 975
DB 717 CCTGAGCAGA-CCTCCGCTGAACCTGTGGACCCCGAGCTCCAGGCCACCCCAAGCGCGCT 775

QY 976 GGCACTGGTGTTCAAACCTCTACCACTTGGAGCCCTGGAGCTGTCTCCCGCCCAACCT 1035
DB 776 GGCACTGGTGTTCAAACCTCTACCACTTGGAGCCCTGGAGCTGTCTCTCCCGCCCAACCT 834

QY 1036 TGTTCCTCATTCGCGCAGACCCATAGCCGCTTGCAGGCGAGAGGACACAGGAGGCCA 1095
DB 835 TGTTCCTCATTCGCGCAGACCCATAGCCGCTTGCAGGCGAGAGGACACAGGAGGCCA 894

QY 1096 GCCTCAGTGCAGACCTTGGGTGGGGGCGCTGGGTCTCTCTCTCCACCCGAGAGGCGACA 1155
DB 895 GCCTCAGTGCAGACCTTGGGTGGGGGCGCATGGGTCTCTCTCTCCACCCGAGCAGC-CA 953

QY 1156 GACACCGGCTTGTGGCAGGCTGGGCTCTGTGTACACCACTCTCTGGGTGGTGCAGAC 1215
DB 954 GACACCGGCTTGTGGCAGGCTGGGCTCTGTGTACACCACTCTCTGGGTGGTGCAGAC 1013

QY 1216 CTTTCCCTCCACCCCGAGGCTTCCAAAGCTCTGTCTCTCTCAGTTTCCAAATGGAACC 1275
DB 1014 CTTTCCCTCCACCCCGAGGCTTCCAAAGCTCTGTCTCTCAGTTTCCAAATGGAACC 1073

QY 1276 ACCTCAGCTCCGAGCAGCCCGACTTACAGGAGCGATGCGCTCTCTCTCTCTCTCTCTCA 1335
DB 1074 AACTCAGCTTCCGAGCAGCCCGAATTACAGGAGCGATGCGCTCTCTCTCTCTCTCTCA 1133

QY 1336 ACCCAGACCCGAGCTCCCTTCTCGCACCCAGAGCTGTGTGGGCGCCAGAGTGTGGGCT 1395
DB 1134 ACCCAGACCCGAGCTCCCTTCTCGCACCCAGAGCTGTGTGGGCGCCAGAGTGTGGGCT 1192

QY 1396 CCGCTCTCTCAGTCTCCAGGCGCTCCGCGCCCAAGTGAAGGGGCGCTCTCTCTCTCTCTCT 1455
DB 1193 CCGCTCTCTCAGTCTCCAGGCGCTCCGCG -CCAAAGTGAAGGGGCGCTCTCTCTCTCTCT 1251

QY 1456 ACACACTGGATTCAGGGCTG-GGGGGGCTTGGCAGATACCTCTCTCTCTCTCTCTCTCTCT 1514
DB 1252 ACACACTGGATTCAGGGCTTGGGGGGCTTGGCAGATACCTCTCTCTCTCTCTCTCTCTCT 1311

QY 1515 AGGCTTTGGGGGCGCTTCCGCGCAGCCCGGGGGCGAGGTAGGGTCTCTGGGGCTTTAGA 1574
DB 1312 AGGCTTTGGGGGCGCTTCCGCGCAG -CCCGGGGGCGAGGTAGGGTCTCTGGGGCTTTAGA 1369

QY 1575 GGCTGGGATGGCTCTCTGGCCCCACCGCAGGGGCAAGCAGCAGGCGGGCTGGGAGGCGG 1634
DB 1370 GGCTGGGATGGCTCTCTGGCCCCACCGCAGGGGCG -AGCGCAGGCGGGCTGGGAGGCGG 1428

QY 1635 CGCGGGCGCTCTGGGCTGGGGGCTCAGTGGAGCGCTG -CTTCCGGGCTGTCTCTCGCGCATC 1693
DB 1429 CGCGGGCGCTCTGGGCTGGGGGCTCAGTGGAGCGCGCCCTCTCGGGGCTGGAGCGCATC 1488

QY 1694 CTTAGTCTCTCGGCGACCCGGGGGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1753
DB 1489 CTTAGTCTCTCGGCGACCCGGGGGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1548

QY 1754 TTTGAGCCAGATCTGTTCATGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1813
DB 1549 TTTGAGCCAGATCTGTTCATGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1608

QY 1814 TTTCTCTGTATGGGGTGGAAAAATCACCGGGAATCCCGCTTTCAGTTCTTTGAAAAAGTTC 1873
DB 1609 TTTCTCTGTATGGGGTGGAAAAATCACCGGGAATCCCGCTTTCAGTTCTTTGAAAAAGTTC 1668

QY 1874 CATGACTCGAATATCTGAAATGAAGAAACAAACCGACTCAACACCTCCAGTAGTCTCC 1933
DB 1669 CATGACTCGAATATCTGAAATGAAGAAACAAACCGACTCAACACCTCCAGTAGTCTCC 1728

QY 1934 AAATGCAATTTTAAATTCGAAAAACAAAAATCTGAAAGAAACGCTCTTTAGTGGCTTTAAG 1993
DB 1729 AAATGCAATTTTAAATTCGAAAAACAAAAATCTGAAAGAAACGCTCTTTAGTGGCTTTAAG 1788

QY 1994 CCCCCAAAACGCTCCCTAAGGCGCTCTCGAGATGAAAGACGGGGGGGAGCCCGCAGCGGTGG 2053
DB 1789 CCCCCAAAACGCTCCCTAAGGCGCTCTCGAGATGAAAGACGGGGGGGAG--CCAGCCAGGTGG 1846

QY 2054 AGACCCCGCAGACCGGGCGCGCGCGGTGACCGAGGCTCCGACAGCGCGCGCGCGCTCTGA 2113
DB 1847 AGACCCCGCAG--CGCGCGCGCGCGCGGTGACCGAGGCTCTGCAAGCGCGCGCGCTGA 1904

QY 2114 GGGTCCGGCGCAGCAGGGTCCAAAGAGGGGCGCGTGTGTCTCGGGTTAAATAGGTT 2173
DB 1905 GGGTCCGGCGCAGCAGGGTCCAAAGAGGGGCGCGTGTGTCTCGGGTTAAATAGGTT 1964

QY 2174 CCGTCCG 2180

Db	1965	CGTCCG 1971	
XX	RESULT 10		
XX	AAC59772		
XX	ID AAC59772 standard; cDNA; 1464 BP.		
XX	AC		
XX	AAC59772;		
DT	26-JAN-2001 (first entry)		
XX	Human secreted protein gene 35 SEQ ID NO:45.		
DE			
XX			
KW	Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;		
KW	antirheumatic; antiproliferative; cytostatic; antibiotic; vasotropic;		
KW	cerebroprotective; neurotropic; neuroprotective; cardiac; virucide;		
KW	fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;		
KW	hyperproliferative disorder; cancer; cardiovascular disorder; infection;		
KW	cerebrovascular disorder; angiogenesis; nervous system disorder;		
KW	ocular disorder; wound healing; skin aging; food additive; preservative;		
OS	ss.		
XX	Homo sapiens.		
XX	WO200056751-A1.		
XX	28-SEP-2000.		
XX	09-MAR-2000; 2000WO-US06013.		
XX	19-MAR-1999; 99US-0125360P.		
PR	11-JUN-1999; 99US-0138626P.		
PR	03-DEC-1999; 99US-0168662P.		
XX			
XX	(HUMA-) HUMAN GENOME SCI INC.		
PI	Rosen CA, Ruben SM, Komatsoulis G;		
XX	WPI; 2000-579482/54.		
DR	P-PSDB; AAB34611.		
XX			
PT	Isolated nucleic acid molecule encoding a human secreted protein is used		
PT	in preventing, treating or ameliorating a medical condition.		
XX			
PS	Claim 1; Page 362; 419pp; English.		
XX			
CC	The polynucleotide sequences given in AAC59738 to AAC59787 encode the		
CC	human secreted proteins given in AAB34577 to AAB34626. AAB34627 to		
CC	AAB34686 represent human secreted polypeptide sequences and proteins		
CC	homologous to them, which are given in the exemplification of the present		
CC	invention. Human secreted proteins have activities based on the tissues		
CC	and cells the genes are expressed in. Example of activities include:		
CC	antiarthritic; immunosuppressive; antirheumatic; antiproliferative;		
CC	cytostatic; cardiac; antiproliferative; antirheumatic; antiproliferative;		
CC	neuroprotective; antirheumatic; vasotropic; cerebroprotective; neurotropic;		
CC	ophthalmological. The polynucleotides and proteins can be used to		
CC	prevent, treat or ameliorate a medical condition in e.g. humans, mice,		
CC	rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used		
CC	in diagnosing a pathological condition or susceptibility to a		
CC	pathological condition. Disorders which are diagnosed or treated include		
CC	autoimmune diseases, hyperproliferative disorders e.g. neoplasms and		
CC	cancers of the breast or liver, cardiovascular disorders, cerebrovascular		
CC	disorders, angiogenesis, nervous system disorders, infections caused by		
CC	bacteria, viruses and fungi and ocular disorders. The proteins can also		
CC	be used to aid wound healing and epithelial cell proliferation, to		
CC	prevent skin aging due to sunburn, to maintain organs before		
CC	transplantation, for supporting cell culture of primary tissues, to		
CC	regenerate tissues and in chemotaxis. The proteins can also be used as a		
CC	food additive or preservative to increase or decrease storage		
CC	capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used		
CC	in the exemplification of the present invention		
XX			

SQ	Sequence 1464 BP; 309 A; 483 C; 408 G; 264 T; 0 U; 0 Other;		
	Query Match 43.1%; Score 938.6; DB 3; Length 1464;		
	Best Local Similarity 95.7%; Pred. No. 4.9e-205;		
	Matches 1040; Conservative 0; Mismatches 39; Indels 8; Gaps 7;		
QY	881 AAGGTCGAGCCCTCAGAGCGGAGCCCGAGAGGCGCTGAGAGAGCTCCCTGAACTG 940		
DB	111 AGGCGCCCGGAGCTCAGAGCGGAGCCCGAGAGGCGCTGAGAGAGCTCCCTGAACTG 170		
QY	941 TGGACCCGAGACTCCGAGCGCCACCCAGGCGCTGGCACTGGTGTCAAACTCCACCA 1000		
DB	171 TGGACCCGAGACTCCGAGCGCCACCCAGGCGCTGGCACTGGTGTCAAACTCCACCA 230		
QY	1001 CTTGAGGCGCTGGAGCTGCTGCCCCCAACCCCTTGTTCATATGCCGAGAGCCATA 1060		
DB	231 CTTGAGGCGCTGGAGCTGCTGT-CCCCCAACCCCTTGTTCATATGCCGAGAGCCATA 289		
QY	1061 GCCCGCTGCAAGGCGAGAGGAGACACAGAGAGCGCCCTGAGTCCGACCTTGGGTGGC 1120		
DB	290 GCCCGCTGCAAGGCGAGAGGAGACACAGAGAGCGCCCTGAGTCCGACCTTGGGTGGC 349		
QY	1121 GGGCGCTGGGTCTCTCGTCCACCCGAGGCGGACAGACCGGCTTGGTGGCAGGCTGG 1180		
DB	350 GGGCGCTGGGTCTCTCGTCCACCCGAGGCGGACAGACCGGCTTGGTGGCAGGCTGG 409		
QY	1181 GCCTCTGTGTCAACCCACTCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1240		
DB	410 GCCTCTGTGTCAACCCACTCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 469		
QY	1241 CCAAGCTCTGCTTCTCAGTTTCCAAATGGAACCACTCACTCCGAGAGCCAGCTTT 1300		
DB	470 CCAAGCTCTGCTTCTCAGTTTCCAAATGGAACCACTCACTCCGAGAGCCAGCTTT 529		
QY	1301 ACCAGAGCGATGCCCTTCCCTCTGCTCCATCAAAACCAAGAGAGCGGACTCCCTTTCT 1360		
DB	530 ACCAGAGCGATGCCCTTCCCTCTGCTCCATCAAAACCAAGAGAGCGGACTCCCTTTCT 589		
QY	1361 GCCACCCGAGGCTGGTCCGCGGCCAGGTGTGGGGTCCGCTCTCTCCACTCCGAGGCTCC 1420		
DB	590 GCCACCCGAGGCTGGTCCGCGGCCAGGTGTGGGGTCCGCTCTCTCCACTCCGAGGCTCC 648		
QY	1421 GCGCCCAAGTGAAGGGGGCCCTTCCGAGAGCTTCAAGAGTTCAGAGGCTTCAAGGCT 1478		
DB	649 GCGCCCAAGTGAAGGGGGCCCTTCCGAGAGCTTCAAGAGTTCAGAGGCTTCAAGGCT 708		
QY	1479 GGGGCTTGGCACATACCTGTCCCTTGGCTGTAGAGAGGCTTGGGGGCGCTTCCGCGGC 1538		
DB	709 GGGGCTTGGCACATACCTGTCCCTTGGCTGTAGAGAGGCTTGGGGGCGCTTCCGCGGC 768		
QY	1539 AGCCCGGGGGCGGAGGTAGGTCTGGGGGCTTAGAGGCTGGGATGGTCTCTGGGCGCCAC 1598		
DB	769 AGCCCGGGGGCGGAGGTAGGTCTGGGGGCTTAGAGGCTGGGATGGTCTCTGGGCGCCAC 827		
QY	1599 GCGCAGGGGGCAAGCGAGCGCGGCTGGAGGCGCGCGCGCGCGCGCTGGGGTGGGGGT 1658		
DB	828 GCGCAGGGGGCG-AGCGAGGCGCGGCTGGAGGCGCGCGCGCGCGCTGGGGTGGGGGT 886		
QY	1659 CAGGTGGAGCTG-CTTCCGGGGCTGGTCCGCGCATCCCTCAGTCCCTGGGCGCCAGCGGGG 1717		
DB	887 CAGGTGGAGCTG-CTTCCGGGGCTGGTCCGCGCATCCCTCAGTCCCTGGGCGCCAGCGGGG 946		
QY	1718 GTGCGTCTCTGTGCGCCACCGCCTGCGGAGCGCTTTTGGACCCAGAGTCTGTTCATGCT 1777		
DB	947 GTGCGTCTCTGTGCGCCACCGCCTGCGGAGCGCTTTTGGACCCAGAGTCTGTTCATGCT 1006		
QY	1778 TTTGCTCTGTGCTGCTGCGGGGGCCCTTTGATGTCTTTCATCTGTAT-GGGGTGGAAAAA 1836		
DB	1007 TTTGCTCTGTGCTGCTGCGGGGGCCCTTTGATGTCTTTCATCTGTATGCGGGTGGAAAAA 1066		
QY	1837 TCACCGGAATCCCGCTTTCAGTCTTTTGAAGAGTTCATGCTCGAATATCTGAAATGA 1896		
DB	1067 TCACCGGAATCCCGCTTTCAGTCTTTTGAAGAGTTCATGCTCGAATATCTGAAATGA 1126		

pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, and polymyositis, rheumatoid arthritis, scleroderma, Sjogren's syndrome, and autoimmune thyroiditis, complications of cancer, infections, and trauma

Query Match	41.9%;	Score 912.8;	DB 2;	Length 1064;
Best Local Similarity	96.1%;	Pred. No. 3.6e-199;		
Matches 957; Conservative	0;	Mismatches 37;	Indels 2;	Gaps 2;

199	QY	TTCTCTGGGGCTCCGGGGCGCGGAGAGCTGATCCCAAGAGAGCGCTTCAGGAGCGGAC	60
1	Db	TTCTCTGGGGCTCCGGGGCGCGGAGAGCTGATCCCAAGAGAGCGCTTCAGGAGCGGAC	60
259	QY	CCGGAGGTGTTTCAAGAGCCAGTGCACAGGACCAAGGGGCCCAAGTCCCAACAGCCATGCA	318
61	Db	CCGGAGGTGTTTCAAGAGCCAGTGCACAGGACCAAGGGGCCCAAGTCCCAACAGCCATGCA	120
319	QY	GACCTGCCCTTGGCAATTCCTTGGCGCAGTTTCCAGGCCCTTGGGACCTTCCTGTTTTT	378
121	Db	GACCTGCCCTTGGCAATTCCTTGGCGCAGTTTCCAGGCCCTTGGGACCTTCCTGTTTTT	180
379	QY	GGCTGCCTCTTGAAGTCTCAGATGAAGGCTGGGACAGCCCCATCTTGCAACAGAGGGGT	438
181	Db	GGCTGCCTCTTGAAGTCTCAGATGAAGGCTGGGACAGCCCCATCTTGCAACAGAGGGGT	240
439	QY	AGTCTCTGTGTCTTGGGCGGAAACAACCGTCAATGTCGTGCAACATCTTCAACGGCTTCTC	498
241	Db	AGTCTCTGTGTCTTGGGCGGAAACAACCGTCAATGTCGTGCAACATCTTCAACGGCTTCTC	300
499	QY	CMATGTCAACATCAAGCTCGTGCCCAAGGACAGAGAGCGGCATCTTCAATAGAGTGGC	558
301	Db	CCATGTCAACATCAAGCTCGTGCCCAAGGACAGAGAGCGGCATCTTCAATAGAGTGGC	360
559	QY	TCCAGGCTACTTCTCCCGGACGGCTGGCAGCTCCAGGTTTCAAGGAGCGCTGGCACAGCT	618
361	Db	TCCAGGCTACTTCTCCCGGACGGCTGGCAGCTCCAGGTTTCAAGGAGCGCTGGCACAGCT	420
619	QY	GGTGATCAAAAGGCGCCCGGACCTCCCATGCTGGGCTGTACATGTGGCACCTCTGTGGACA	678
421	Db	GGTGATCAAAAGGCGCCCGGACCTCCCATGCTGGGCTGTACATGTGGCACCTCTGTGGACA	480
679	QY	CCAGAGAAATAACAGACAAAGTCAACGCTGGAGGTTTCAAGTGACAGAACCCCAAGTCGCGCC	738
481	Db	CCAGAGAAATAACAGACAAAGTCAACGCTGGAGGTTTCAAGTGACAGAACCCCAAGTCGCGCC	540
739	QY	TGACACTGGGTTCTGGCCTGTGCGACGGTGTGCTACTGTCTTCACTCTCTTGGTGGC	798
541	Db	CGACACTGGGTTCTGGCCTGTGCGACGGTGTGCTACTGTCTTCACTCTCTTGGTGGC	600
799	QY	TCGTGTCATGTTTCGCTGGTTACAGGTGCGCTGTTCCCAAGCAACGCGGGAGAAAGTT	858
601	Db	TCGTGTCATGTTTCGCTGGTTACAGGTGCGCTGTTCCCAAGCAACGCGGGAGAAAGTT	660
859	QY	CTTCTCTTAGAACCCCAAGTTCGAGCGCTTCAGAGCGGAGCCAGCAGGGCCT	918
661	Db	CTTCTCTTAGAACCCCAAGTTCGAGCGCTTCAGAGCGGAGCCAGCAGGGCCT	720
919	QY	GAGCAGAGCTTCGGTGAATGTGTGACCCAGACTCCGAGCCACCCAGGCGCGTGGC	978
721	Db	GAGCAGAGCTTCGGTGAATGTGTGACCCAGACTCCGAGCCACCCAGGCGCGTGGC	780
979	QY	ACTGGTGTCAAAACCTTCACACTTGGAGCGCTTGGAGCTGTCTGTCTCCCGCCCAACCTTGT	1038
781	Db	ACTGGTGTCAAAACCTTCACACTTGGAGCGCTTGGAGCTGTCTGTCTCCCGCCCAACCTTGT	839
1039	QY	TTCCATATGCGCAGACCCATAGCCGCTTCGAAGGACAGAGACACAGGAGCGCAGCC	1098
840	Db	TTCCATATGCGCAGACCCATAGCCGCTTCGAAGGAGAGAGACACAGGAGTAGCCACC	899
1099	QY	CTGAGTCCGACCTTGGGTGGC-GGGGCTGGGTCTCTGTCTCCCAACCGGAGGGCACAGA	1157
900	Db	CTGAGTCCGACCTTGGGTGGGCGGGGCTTGGGTCTCTGTCTCCCAACCGGAGGGCACAGA	959

QY 1158 CACCGGCTTGCTGGCAGCTGGGCTCTGTGTTCAC 1193
Db | | | | | | | | | | | | | | | | | | | | | |
960 AGACACCGGCTTGCTGGCAGGCTTGCGGGCTC 995

RESULT 12

ADB36352
ID ADB36352 standard; cDNA; 958 BP.

AC ADB36352;

XX 04-DEC-2003 (first entry)

DE Human immune response associated protein IRAP-7 cDNA SEQ ID NO:23.

XX human; immune response associated protein; IRAP; anti-HIV; antiallergic;
KW antianemic; antiasthmatic; antiarteriosclerotic; dermatological;
KW antiinflammatory; antidiabetic; nephrotropic; antichyroid; thyromimetic;
KW immunosuppressive; antirheumatic; antiarthritic; osteopathic;
KW antibacterial; virucide; antiparasitic; protozoacide; fungicide;
KW cerebroprotective; neuroprotective; nootropic; antiparkinsonian;
KW antipsoriatic; cytostatic; cardiant; gene therapy;
KW immune system disorder; neurological disorder; developmental disorder;
KW muscle disorder; cell proliferative disorder; AIDS; allergy; anaemia;
KW asthma; atherosclerosis; contact dermatitis; Crohn's disease; diabetes;
KW glomerulonephritis; Grave's disease; Hashimoto's thyroiditis;
KW multiple sclerosis; rheumatoid arthritis; osteoporosis;
KW systemic lupus erythematosus; infection; stroke; Alzheimer's disease;
KW Parkinson's disease; psoriasis; cancer; cardiomyopathy; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
FH 116..565
CDS /tag= a
FT /product= "IRAP-7"

PN WO2003074726-A2.

XX 12-SEP-2003.

PF 28-FEB-2003; 2003WO-US006307.

XX 01-MAR-2002; 2002US-0361088P.

PR 27-MAR-2002; 2002US-0368494P.

PR 10-MAY-2002; 2002US-0379876P.

PR 28-JUN-2002; 2002US-0392641P.

XX (INCY-) INCYTE GENOMICS INC.

XX Chawla NK, Becha SD, Jin P, Tran UK, Kable AE, Lee S;

PI Marquis JP, Emerling BM, Ison CH, Ramkumar J, Chang H;

XX WPI; 2003-722079/68.

DR P-PSDB; ADB36336.

XX New human immune response associated proteins and polynucleotides, useful

PT for diagnosing, preventing or treating diseases or conditions associated

PT with aberrant enzyme expression, e.g. cancer, AIDS, atherosclerosis or

PT stroke.

XX Claim 5; Page 152; 158pp; English.

XX ADB36346 to ADB36361 encode the human immune response associated proteins

CC given in ADB36330 to ADB36345, designated IRAP-1 to IRAP-16. IRAP

CC sequences have anti-HIV, antiallergic, antianemic, antiasthmatic,

CC antiarteriosclerotic, dermatological, antiinflammatory, antidiabetic,

CC proteins and polynucleotides can be used in diagnosing, preventing or

CC treating diseases or conditions associated with the decreased expression

CC or overexpression of IRAP, such as immune system, neurological,

CC developmental, muscle or cell proliferative disorders. The disorders may

CC include AIDS, allergies, anaemia, asthma, atherosclerosis, contact

CC dermatitis, Crohn's disease, diabetes, glomerulonephritis, Grave's

CC disease, Hashimoto's thyroiditis, multiple sclerosis, rheumatoid

CC arthritis, osteoporosis, systemic lupus erythematosus, infections (e.g.

CC bacterial, viral, fungal, parasitic or protozoal), stroke, Alzheimer's

CC disease, Parkinson's disease, psoriasis, cancer and cardiomyopathy. They

CC are also useful in assessing the effects of exogenous compounds on the

CC expression of nucleic acid and amino acid sequences of IRAP. The IRAP or

CC its fragments are useful in screening compounds for effectiveness as

CC agonist or antagonist of the polypeptides, or in altering the expression

CC of the target polynucleotide and compounds that specifically bind to or

CC modulate the activity of the polypeptide.

XX Sequence 958 BP; 186 A; 313 C; 285 G; 174 T; 0 U; 0 Other;

SQ Query Match 31.8%; Score 692.4; DB 10; Length 958;

Best Local Similarity 85.5%; Pred. No. 1.2e-148;

Matches 852; Conservative 0; Mismatches 36; Indels 108; Gaps 3;

QY 199 TTCCTGGGGCTCCGGGGCGCGAGAGCTGCATCCACAGAGAGCGCTCCAGAGCGGAC 258

Db 1 TTCTCTGGGGCTCCGGGGCGCGAGAGCTGCATCCACAGAGAGCGCTCCAGAGCGGAC 60

QY 259 CCGGAGAGTGTTCAGAGCCAGTAGTCAAGAGACAGAGGCGCCCAAGTCCACACCATGCA 318

Db 61 CCGGAGAGTGTTCAGAGCCAGTAGTCAAGAGACAGAGGCGCCCAAGTCCACACCATGCA 120

QY 319 GACCTGCCCCCTGGCATTCCTCTGGGCGACGCTTTCCAGAGCGCTTGGGAGCCCTCTCTGTTT 378

Db 121 GACCTGCCCCCTGGCATTCCTCTGGGCGACGCTTTCCAGAGCGCTTGGGAGCCCTCTCTGTTT 180

QY 379 GGCTGCTCTTGTAGTGTCTCAGATGAAGGCTGGGACAGCCCATCTTGCACAGAGGGGT 438

Db 181 GGCTGCTCTTGTAGTGTCTCAGATGAAGGCTGGGACAGCCCATCTTGCACAGAGGGGT 240

QY 439 AGTCTCTGTGTTTGGGGCGAGACACCGTGCATGCTCTGCAACATCTCCAAGCGCTTCTC 498

Db 241 AGTCTCTGTGTTTGGGGCGAGACACCGTGCATGCTCTGCAACATCTCCAAGCGCTTCTC 300

QY 499 CCATGTCAACATCAAGCTGCGTGGCCACAGGGCAGAGAGCGCCATCTTCAATGAGGTGGC 558

Db 301 CCATGTCAACATCAAGCTGCGTGGCCACAGGGCAGAGAGCGCCATCTTCAATGAGGTGGC 360

QY 559 TCCAGGCTACTTCTCCGGGACGGCTGGCAGCTCCAGGTTTCAAGGAGGGGTGGCAGCT 618

Db 361 TCCAGGCTACTTCTCCGGGACGGCTGGCAGCTCCAGGTTTCAAGGAGGGGTGGCAGCT 420

QY 619 GGTGATCAAGGCGCCCGGGACTCCCATGCTGGGCTGTACATGTGCGACCTCTGTTGGACA 678

Db 421 GGTGATCAAGGCGCCCGGGACTCCCATGCTGGGCTGTACATGTGCGACCTCTGTTGGACA 480

QY 679 CCAGAGAAATAACAGACAAAGTCAAGCTGGAGGTTTTCAGGTGCGAGAAACCCAGTCCGCCCC 738

Db 481 CCAGAGAAATAACAGACAAAGTCAAGCTGGAGGTTTTCAGGTGCGAGAAACCCAGTCCGCCCC 515

QY 739 TGACACTGGGTTTGGGCTGTGCGCAGCGGTGGTCACTGCTGTCTTCATCTCTTGGTGGC 798

Db 516 ----- 515

QY 799 TCTGTCATGTTTCGCTGTACAGTGGCTGTGTTTCCAGCAACCGCGGAGAAAGT 858

Db 516 -----CAGTGGCTGTGTTTCCAGCAACCGCGGAGAAAGT 554

QY 859 CTTCTCTAGAACCCAGATCAAGGTTCGAGCCCTCAGAGCGGAGGCCACAGAGGCT 918

Db 555 CTTCTCTAGAACCCAGATCAAGGTTCGAGCCCTCAGAGCGGAGGCCACAGAGGCT 614

QY 919 GAGCAGAGCTTCCGCTGAACCTGTGGACCCCAAGACTCCAGGCCACCCCAAGGCCGCTGGC 978

Db 615 GAGCAGAGCCTCGGTGAACCTGTGAGACCCCGAGACTCCGAGCCCAAGCCGCTGGC 674

QY 979 ACTGGTGTTCACACCTCACCCTTGGAGCCCTGGAGCTGTGTGCTCCCGCCCAACCTTGT 1038

Db 675 ACTGGTGTTCACACCTCACCCTTGGAGCCCTGGAGCTGTGTGCTCCCGCCCAACCTTGT 733

QY 1039 TTCCATATGTCGCGCAGACCCATAGCCGCTGCAAGCGCAGAGGACACAGGAGGCCAGCC 1098

Db 734 TTCCATATGTCGCGCAGACCCATAGCCGCTGCAAGCGCAGAGGACACAGGAGGTAGCCACC 793

QY 1099 CTGAGTCCGACCTTGGGTGGC-GGGGCTGGGTCTCTGCTCCACCCGAGGGGCACAGA 1157

Db 794 CTGAGTCCGACCTTGGGTGGC-GGGGCTGGGTCTCTGCTCCACCCGAGGGGCACAGA 853

QY 1158 CACCGGCTTGTGGCAGGCTGGGCTCTGTGTCTAC 1193

Db 854 AGNACCGGCTTGTGGCAGGCTTGGGGCTC 889

RESULT 13

ID ADB36353 standard; cDNA; 942 BP.

AC ADB36353;

XX 04-DEC-2003 (first entry)

XX Human immune response associated protein IRAP-8 cDNA SEQ ID NO:24.

XX human; immune response associated protein; IRAP; anti-HIV; antiallergic;

KW antianemic; antidiabetic; antiarteriosclerotic; dermatological;

KW antiinflammatory; antidiabetic; nephrotropic; antichyroid; thyromimetic;

KW immunosuppressive; antirheumatic; antiarthritic; osteopathic;

KW antibacterial; viricide; antiparasitic; protozoacide; fungicide;

KW cerebroprotective; neuroprotective; nootropic; antiparkinsonian;

KW antipsoriatic; cytostatic; cardiant; gene therapy;

KW immune system disorder; neurological disorder; developmental disorder;

KW muscle disorder; cell proliferative disorder; AIDS; allergy; anaemia;

KW asthma; atherosclerosis; contact dermatitis; Crohn's disease; diabetes;

KW glomerulonephritis; Grave's disease; Hashimoto's thyroiditis;

KW multiple sclerosis; rheumatoid arthritis; osteoporosis;

KW systemic lupus erythematosus; infection; stroke; Alzheimer's disease;

KW Parkinson's disease; psoriasis; cancer; cardiomyopathy; gene; ss.

XX Homo sapiens.

OS

XX

FH Key Location/Qualifiers

FT CDS 101..232

FT /*tag= a

FT /product= "IRAP-8"

FT

XX WO2003074726-A2.

XX 12-SEP-2003.

XX 28-FEB-2003; 2003WO-US006307.

XX 01-MAR-2002; 2002US-0361088P.

XX 27-MAR-2002; 2002US-0368494P.

XX 10-MAY-2002; 2002US-0379876P.

XX 28-JUN-2002; 2002US-0392641P.

XX (INCY-) INCYTE GENOMICS INC.

XX Chawla NK, Becha SD, Jin P, Tran UK, Kable AE, Lee S;

PI Marquis JP, Emerling BW, Ison CH, Ramkumar J, Chang H;

XX

DR WPI; 2003-722079/68.

DR P-PSDB; ADB36337.

XX New human immune response associated proteins and polynucleotides, useful

PT for diagnosing, preventing or treating diseases or conditions associated

PT with aberrant enzyme expression, e.g. cancer, AIDS, atherosclerosis or

stroke.

Claim 5; Page 153; 158pp; English.

ADB36346 to ADB36361 encode the human immune response associated proteins

given in ADB36330 to ADB36345, designated IRAP-1 to IRAP-16. IRAP

sequences have anti-HIV, antiallergic, antianemic, antidiabetic,

antiarteriosclerotic, dermatological, antinflammatory, antidiabetic,

nephrotropic, antithyroid, thyromimetic, immunosuppressive,

antirheumatic, antiarthritic, osteopathic, antibacterial, viricide,

antipsoriatic, protozoacide, fungicide, cerebroprotective,

neuroprotective, nootropic, antiparkinsonian, antichyroid, cytostatic

and cardiant activities, and can be used in gene therapy. The IRAP

proteins and polynucleotides can be used in diagnosing, preventing or

treating diseases or conditions associated with the decreased expression

or overexpression of IRAP, such as immune system, neurological,

developmental, muscle or cell proliferative disorders. The disorders may

include AIDS, allergies, anaemia, asthma, atherosclerosis, contact

dermatitis, Crohn's disease, diabetes, glomerulonephritis, Grave's

disease, Hashimoto's thyroiditis, multiple sclerosis, rheumatoid

arthritis, osteoporosis, systemic lupus erythematosus, infections (e.g.

bacterial, viral, fungal, parasitic or protozoal), stroke, Alzheimer's

disease, Parkinson's disease, psoriasis, cancer and cardiomyopathy. They

are also useful in assessing the effects of exogenous compounds on the

expression of nucleic acid and amino acid sequences of IRAP. The IRAP or

its fragments are useful in screening compounds for effectiveness as

agonist or antagonist of the polypeptides, or in altering the expression

of the target polynucleotide and compounds that specifically bind to or

modulate the activity of the polypeptide.

Sequence 942 BP; 175 A; 310 C; 282 G; 175 T; 0 U; 0 Other;

Query Match 30.9%; Score 673.8; DB 10; Length 942;

Best Local Similarity 85.1%; Pred. No. 2.3e-144;

Matches 835; Conservative 0; Mismatches 37; Indels 109; Gaps 3;

QY 214 GCGCGGAGAGCTGCATCCAGAGGAGCGGTCCAGAGCGGACCCCGGAGTGTTCAA 273

Db 1 GCGCGGAGAGCTGCATCCAGAGGAGCGGTCCAGAGCGGACCCCGGAGTGTTCAA 60

QY 274 GAGCAGTGAACAAGGACAGAGGCGGCCCAAGTCCACAGGACCATGAGCCCTGGC 333

Db 61 GAGCAGTGAACAAGGACAGAGGCGGCCCAAGTCCACAGGACCATGAGCCCTGGC 120

QY 334 ATTCCCTGGCAGCGTTTCCAGAGCGCTTGGAGACCTCTGTTTGGTGCCTCTTCAG 393

Db 121 ATTCCCTGGCAGCGTTTCCAGAGCGCTTGGAGACCTCTGTTTGGTGCCTCTTCAG 180

QY 394 TGCTCAGAAATGAAGGCTGGGACAGCCCATCTGCACAGAGGGGGTAGTCTCTGTGTCTTG 453

Db 181 TGCTCAGAAATG----- 191

QY 454 GGGCGAGAAACACCGTCATGTCTGTCGCAACATCTCCAAAGCCTTCTCCCATGTCAACATCAA 513

Db 192 -----AA 193

QY 514 GCTGGTGGCCACGGGCGAGAGCGCCATCTTCAATGAGGTGGCTCCAGGCTACTTCTC 573

Db 194 GCTGGTGGCCACGGGCGAGAGCGCCATCTTCAATGAGGTGGCTCCAGGCTACTTCTC 253

QY 574 CCGGAGACGGCTGGCAGCTCCAGGTTTCCAGGAGCGGTGGCAGCAGTGTGTGATCAAGAGGCGC 633

Db 254 CCGGAGACGGCTGGCAGCTCCAGGTTTCCAGGAGCGGTGGCAGCAGTGTGTGATCAAGAGGCGC 313

QY 634 CCGGAGCTCCCATGTGGGCTGTACATGTGGGACCTCGTGGGACACCAAGAGAAATAACAG 693

Db 314 CCGGAGCTCCCATGTGGGCTGTACATGTGGGACCTCGTGGGACACCAAGAGAAATAACAG 373

QY 694 ACAAGTCACGTGGAGGTTTTCAGGTGCAGAAACCCAGTCCGCCCTTCAGCATCTGGGTTCTG 753

Db 374 ACAAGTCACGTGGAGGTTTTCAGGTGCAGAAACCCAGTCCGCCCTTCAGCATCTGGGTTCTG 433

QY 754 GCCTGTGCCAGAGCGGTGGTCACTGTCTTCTTCTTGGTGCCTCTGTGTCTGTTCGC 813

Db 517 ----- 516
QY 799 TCTGTCATGTTGCGCTGTACAGTGCCTGTTCCAGCAACGCCGGGAGAGAGTT 858
Db 517 -----AGAGAGTT 526
QY 859 CTTCTCTCTAGAACCCAGATGAAGTCCGAGCCCTCAGAGCGGAGCCAGCAGGCGCT 918
Db 527 CTTCTCTCTAGAACCCAGATGAAGTCCGAGCCCTCAGAGCGGAGCCAGCAGGCGCT 586
QY 919 GAGCAGAGCCTCCGCTGAAGTGTGACCCAGACTCCGAGCCCAAGCGGCGCTGGC 978
Db 587 GAGCAGAGCCTCCGCTGAAGTGTGACCCAGACTCCGAGCCCAAGCGGCGCTGGC 646
QY 979 ACTGCTGTCAACCCCTCACCATTGGAGCCCTGGAGCTGTGTCTCCCGCCCAACCTTTGT 1038
Db 647 ACTGCTGTCAACCCCTCACCATTGGAGCCCTGGAGCTGTGTCTCCCGCCCAACCTTTGT 705
QY 1039 TTCCATATGCCGAGACCCATAGCCGCTGCAAGCGCAGAGGACACAGGAGCCAGCC 1098
Db 706 TTCCATATGCCGAGACCCATAGCCGCTGCAAGCGCAGAGGACACAGGAGGTAGCCACC 765
QY 1099 CTGAGTCCGACTTGGTGCC-GGGGCTGGGTCTCTGTCCTCCACCGAGGCGACAGA 1157
Db 766 CTGAGTCCGACTTGGTGCCGGGGGCTGGGTCTCTGTCCTCCACCGAGGCGACAGA 825
QY 1158 CACCGCTTGTCTGGCAGGCTGGGCTCTGTGTAC 1193
Db 826 AGACACGGGCTTGTCTGGCAGGCTTGGGGCTC 861

RESULT 15

ID AAL60550
AL60550 standard; cDNA; 869 BP.

AC AAL60550;

DT 03-SEP-2003 (first entry)

DE Human organelle-associated protein (ORGA)-10 cDNA.

KW Human; organelle-associated protein; ORGA; cell proliferative disorder;
KW actinic keratosis; arteriosclerosis; reproductive disorder; infertility;
KW endometriosis; gastrointestinal disorder; ectopic pregnancy; hepatitis;
KW gastroenteritis; neurological disorder; Alzheimer's disease; cirrhosis;
KW glomerulonephritis; renal amyloidosis; renal failure; Addison's disease;
KW renal disorder; ovulatory defect; teratogenesis; Pick's disease; cancer;
KW Huntington's disease; urologic disorder; cystitis; dementia; dysphagia;
KW indigestion; gastritis; hypertension; gene therapy; bursitis; anorexia;
KW infection; urethritis; prostatitis; gene; ss.

XX Homo sapiens.

OS Location/Qualifiers

PH Key 116..667

FT CDS /*tag= a

FT /*product= "Human ORGA protein"

XX WO2003044171-A2.

XX 30-MAY-2003.

XX 15-NOV-2002; 2002WO-US036807.

XX 16-NOV-2001; 2001US-0332384P.

XX 13-DEC-2001; 2001US-0341187P.

XX 23-JAN-2002; 2002US-0351151P.

XX 27-FEB-2002; 2002US-0360269P.

XX 05-APR-2002; 2002US-0370637P.

XX 14-JUN-2002; 2002US-0388946P.

XX (INCY-) INCYTE GENOMICS INC.

XX

PI Ramkumar J, Burford N, Xu Y, Chawla NK, Kable AE, Becha SD;

PI Richardson TW, Gervad AE, Sprague WM, Elliott VS, Khare R;

PI Hawkins PR, Jin P, Lee SY, Marquis JP, Swarnakar A, Wilson AD;

PI Hafalia AJA, Bulloch S;

XX WPI; 2003-457603/43.

DR P-PSDB; AAO29897.

XX New organelle-associated proteins and polynucleotides, useful for

PT diagnosing, treating and/or preventing cell proliferative, reproductive,

PT gastrointestinal, neurological, urologic, and renal disorders.

XX Claim 5; Page 184-185; 194pp; English.

XX The invention relates to human organelle-associated proteins (ORGA) and

CC nucleic acid molecules encoding such proteins. ORGA sequences are useful

CC for diagnosing, treating and/or preventing cell proliferative disorders

CC (e.g. actinic keratosis, arteriosclerosis, bursitis, cirrhosis, cancers

CC or hepatitis), reproductive disorders (e.g. infertility, endometriosis,

CC ovulatory defects, ectopic pregnancy or teratogenesis), gastrointestinal,

CC disorders (e.g. dysphagia, peptidic oesophagitis, gastritis, indigestion,

CC anorexia or gastroenteritis), neurological disorders (e.g. Alzheimer's

CC disease, Pick's disease, Huntington's disease or dementia), urologic

CC disorders (e.g. infection, urethritis, cystitis, prostatitis or upper

CC tract infections) and renal disorders (e.g. glomerulonephritis, renal

CC amyloidosis, renal failure, Addison's disease or hypertension). The

CC invention is also useful in gene therapy. The present sequence is human

CC ORGA protein encoding cDNA

XX Sequence 869 BP; 161 A; 289 C; 262 G; 157 T; 0 U; 0 Other;

SQ Query Match 23.5%; Score 512.8; DB 9; Length 869;

Best Local Similarity 76.5%; Pred. No. 1.9e-107;

Matches 762; Conservative 0; Mismatches 37; Indels 197; Gaps 3;

QY 199 TTCTGGGGCTCCGGGGCGGAGAACTGCATCCCAAGAGAGCGCTCCAGAGCGGAC 258

Db 1 TTCTGGGGCTCCGGGGCGGAGAACTGCATCCCAAGAGAGCGCTCCAGAGCGGAC 60

QY 259 CCGGAGTGTTCAGAGAGCCAGTCAAGAGACAGGGGGCCCAAGTCCACAGCCATGCA 318

Db 61 CCGGAGTGTTCAGAGAGCCAGTCAAGAGACAGGGGGCCCAAGTCCACAGCCATGCA 120

QY 319 GACCTGCCCTCCGATTCCTGCGCCAGTTCCTCCAGGCCCTTCGACCCCTCTCTTTT 378

Db 121 GACCTGCCCTCCGATTCCTGCGCCAGTTCCTCCAGGCCCTTCGACCCCTCTCTTTT 180

QY 379 GGCTCCCTCTTGTAGTGTCTCAGAAATGAAGGCTGGGACAGCCCATCTTGCACAGAGGGGT 438

Db 181 GGCTCCCTCTTGTAGTGTCTCAGAAATGAAGGCTGGGACAGCCCATCTTGCACAGAGGGGT 240

QY 439 AGTCTCTGTCTTGGGGCGGAGAACCCGTCTATGTCCTGCAACATCTCCAAGCCCTCTC 498

Db 241 AGTCTCTGTCTTGGGGCGGAGAACCCGTCTATGTCCTGCAACATCTCCAAGCCCTCTC 300

QY 499 CCATGTCAACATCAAGCTCGTGCCCAAGGAGAGAGCGCCATCTTCAATAGGTTGC 558

Db 301 CCATGTCAACATCAAGCTCGTGCCCAAGGAGAGAGCGCCATCTTCAATAGGTTGC 360

QY 559 TCCAGGCTACTTCTCCCGGACGGGTGCGAGTCCAGTTTCAGGAGCGGTGGCAGAGCT 618

Db 361 TCCAGGCTACTTCTCCCGGACGGGTGCGAGTCCAGTTTCAGGAGCGGTGGCAGAGCT 420

QY 619 GGTGATCAAGGGCGCGGAGCTCCCATGCTGGGTGTACATGTGGCAGCTCTCGTGGACA 678

Db 421 GGTGATCAAGGGCGCGGAGCTCCCATGCTGGGTGTACATGTGGCAGCTCTCGTGGACA 480

QY 679 CCAGAGAAATACAGACAAGTCAAGTTCAGGTGGAGTTTCAGGTGAGAACCCAGCTCCGCC 738

Db 481 ----- 480

QY 739 TGACACTGGGTTCTGGCCTGTGCCAGCGGTGGTCACTGCTGTCTTCTCATCTCTTGTGTCG 798

```
Db 481 ----- 480
Qy 799 TCTGTCATGTTTCGCTGTGTACAGTGCCTGTTCCAGCAACGCGGGAGAGAAGTT 858
Db 481 ----- 480
Qy 859 TTTCCTCTAGAACCCCAAGTGCAGCCCTCAGAGCGGAGCCAGCAGGGCCT 918
Db 481 ----- 525
Qy 919 GAGCAGAGCCTCGCTGAACTGTGACCCCACTCCGAGCCCAACGCGCGTGGC 978
Db 526 GAGCAGAGCCTCGCTGAACTGTGACCCCACTCCGAGCCCAACGCGCGTGGC 585
Qy 979 ACTGCTGTTCAAACCTCACCATTGGAGCCCTGGAGCTGCTGTCCCTCCCAACCTTGT 1038
Db 586 ACTGCTGTTCAAACCTCACCATTGGAGCCCTGGAGCTGCTGT-CCCCCAACCTTGT 644
Qy 1039 TTCCATATGCCCGCAGACCCATAGCCGCTGCAAGGAGAGAGGACACAGGAGAGCCAGCC 1098
Db 645 TTCCATATGCCCGCAGACCCATAGCCGCTGCAAGGAGAGAGGACACAGGAGTAGCCACC 704
Qy 1099 CTGAGTCCGACCTTGGGTGGC-GGGGCTGGGTCTCTGTCCTCCACCGGAGGACACAGA 1157
Db 705 CTGAGTCCGACCTTGGGTGGGCGGGGCTGGGTCTCTGTCCTCCACCGGAGGACACA 764
Qy 1158 CACCGGCTTGTGGCAGGCTGGGCTCTGTGTAC 1193
Db 765 AGACACCGGGCTTTGCTTGGCAAGCTTGGGCGCTC 800
```

Search completed: March 20, 2005, 02:10:42

Job time : 1195 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 01:40:02 ; Search time 380 Seconds
(without alignments)

9387.061 Million cell updates/sec

Title: US-10-080-522-2

Perfect score: 2180

Sequence: 1 ATTCCTGCTTCCTTAGCGT.....TTAAATAAGTTCGTCG 2180

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2180	100.0	2180	2	US-08-755-559-2
2	2180	100.0	2180	3	US-09-210-474-2
3	2180	100.0	2180	3	US-09-539-774-2
4	1951	89.5	2000	4	US-09-997-165-3
5	1951	89.5	2000	4	US-09-949-016-426
6	1889.6	86.7	2001	4	US-09-949-016-2916
7	1241.6	57.0	16738	4	US-09-949-016-12168
8	1241.6	57.0	16738	4	US-09-949-016-14678
9	566	26.0	601	4	US-09-949-016-26333
10	566	26.0	601	4	US-09-949-016-26333
11	541.4	24.8	601	4	US-09-949-016-105202
12	541.4	24.8	601	4	US-09-949-016-26334
13	155.4	7.1	1112	4	US-09-949-016-105203
14	138.6	6.4	601	4	US-09-949-016-26332
15	138.6	6.4	601	4	US-09-949-016-105201
16	55	2.5	7218	1	US-08-232-463-14
17	48.4	2.2	19203	4	US-09-949-016-15519
18	48.4	2.2	19210	4	US-09-949-016-12129
19	48	2.2	31467	4	US-09-949-016-13134
20	48	2.2	31868	4	US-09-949-016-11907
21	46.2	2.1	4112	1	US-08-340-203A-2
22	46.2	2.1	4112	2	US-08-452-567-2
23	46.2	2.1	4112	3	US-08-452-567-2
24	46.2	2.1	4112	3	US-09-085-407-2
25	46.2	2.1	4616	1	US-08-340-203A-1
26	46.2	2.1	4616	2	US-08-452-567-1
27	46.2	2.1	4616	2	US-08-452-427-1

c 28	46.2	2.1	4616	3	US-09-085-407-1	Sequence 1, Appli
c 29	45.8	2.1	61158	4	US-09-949-016-15041	Sequence 15041, A
c 30	45.8	2.1	75431	4	US-09-949-016-15122	Sequence 15122, A
c 31	45.8	2.1	670689	4	US-09-949-016-12505	Sequence 12505, A
c 32	45.8	2.1	670690	4	US-09-949-016-14207	Sequence 14207, A
c 33	45.4	2.1	3561	4	US-09-616-289-48	Sequence 48, Appl
c 34	45.2	2.1	3120	4	US-09-016-434-1050	Sequence 1050, Ap
c 35	45	2.1	46244	4	US-09-949-016-13508	Sequence 13508, A
c 36	44.8	2.1	6029	4	US-09-949-016-13342	Sequence 13342, A
c 37	44.8	2.1	11093	4	US-09-949-016-12142	Sequence 12142, A
c 38	44.6	2.0	1746	4	US-09-573-080A-186	Sequence 186, App
c 39	44.6	2.0	152331	3	US-09-128-155-16	Sequence 16, Appl
c 40	44.2	2.0	154746	4	US-09-827-688-8	Sequence 8, Appli
c 41	44.2	2.0	154746	4	US-09-827-688-8	Sequence 8, Appli
c 42	44	2.0	2823	1	US-08-398-008A-1	Sequence 1, Appli
c 43	44	2.0	2823	2	US-08-893-333-1	Sequence 1, Appli
c 44	44	2.0	31391	4	US-09-949-016-14319	Sequence 14319, A
c 45	43.8	2.0	430	4	US-09-621-976-16556	Sequence 16556, A

ALIGNMENTS

RESULT 1
US-08-755-559-2
; Sequence 2, Application US/08755559
; Patent No. 5912142
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; NUMBER OF SEQUENCES: CELLS
; CORRESPONDENCE ADDRESS: 2
; ADDRESSEE: NIXON & VANDERHVE P. C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,559
; FILING DATE: 22-NOV-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2180 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-755-559-2

Query Match 100.0%; Score 2180; DB 2; Length 2180;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTCCTGCTTCCTTAGCGTGAACCGGGTGGCGCTCCCGTGAATAATAATTAC 60
DB 1 ATTCCTGCTTCCTTAGCGTGAACCGGGTGGCGCTCCCGTGAATAATAATTAC 60

QY	61	CGTACGCTTGTGTGAACGGGGTGGTTCCGAAACTTGGAGGCTTCCCGTAAACCCAG	120	1141	CACCCGAGGGCAGACACACCGGCTGCTTGGCAGGCTGGGCTCTGTGTCAACCCATCC	1200
Db	61	CGTACGCTTGTGTGAACGGGGTGGTTCCGAAACTTGGAGGCTTCCCGTAAACCCAG	120	1201	TGGGTGCGTGCAGACCCCTTCCCTCCCTCCAGCCCGCAGGTCTTCCAGCTCTGCTTCTCACT	1260
QY	121	CTCTCTTCTCACTCGGAGGTGGTCCCGGGGGTCCCGCCCTCTCTCTCTCTCTCTCTCT	180	1201	TGGGTGCGTGCAGACCCCTTCCCTCCCTCCAGCCCGCAGGTCTTCCAGCTCTGCTTCTCACT	1260
Db	121	CTCTCTTCTCACTCGGAGGTGGTCCCGGGGGTCCCGCCCTCTCTCTCTCTCTCTCTCT	180	1261	TTCCAAATGGAAACCACTTCCCTCCCTCCAGCCCGCAGGTCTTCCAGCTCTGCTTCTCACT	1320
QY	181	CTCTCTGCTCTTCTCACTTCTCTGCTGCTTCTCTGCTGCTTCTCTGCTGCTTCTCTGCT	240	1261	TTCCAAATGGAAACCACTTCCCTCCCTCCAGCCCGCAGGTCTTCCAGCTCTGCTTCTCACT	1320
Db	181	CTCTCTGCTCTTCTCACTTCTCTGCTGCTTCTCTGCTGCTTCTCTGCTGCTTCTCTGCT	240	1321	CTCTGCTCTCATCAAAACCCACAGACCCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1380
QY	241	GGCGGTCCAGAGCGGACCCCGGAGTGTTCAGAGCCAGTGACAAAGACCAAGGGGCCCA	300	1321	CTCTGCTCTCATCAAAACCCACAGACCCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1380
Db	241	GGCGGTCCAGAGCGGACCCCGGAGTGTTCAGAGCCAGTGACAAAGACCAAGGGGCCCA	300	1381	CCCGAGGTGGGGTCCGCT	1440
QY	301	AGTCCACAGCCATGACAGCTCCCTCTGCTGCTTCTCTGCTGCTTCTCTGCTGCTTCTCT	360	1381	CCCGAGGTGGGGTCCGCT	1440
Db	301	AGTCCACAGCCATGACAGCTCCCTCTGCTGCTTCTCTGCTGCTTCTCTGCTGCTTCTCT	360	1441	CTGCGGAGGCTCAGACACACTGAGTTCAGAGGCTGGGGGGCTTGGGCAATACCTGTC	1500
QY	361	TGGGACCTCTCTGCTTCTGCTGCTTCTCTGCTGCTTCTCTGCTGCTTCTCTGCTGCTTCT	420	1441	CTGCGGAGGCTCAGACACACTGAGTTCAGAGGCTGGGGGGCTTGGGCAATACCTGTC	1500
Db	361	TGGGACCTCTCTGCTTCTGCTGCTTCTCTGCTGCTTCTCTGCTGCTTCTCTGCTGCTTCT	420	1501	CTTGGCTATGAGCAGGCTTGGGGGCTTCCCGGCGAGCCCGGGGGGCGGAGGTAGGG	1560
QY	421	CATCTGCACAGAGGGGTAGTCTCTGCTGCTTCTCTGCTGCTTCTCTGCTGCTTCTCTGCT	480	1501	CTTGGCTATGAGCAGGCTTGGGGGCTTCCCGGCGAGCCCGGGGGGCGGAGGTAGGG	1560
Db	421	CATCTGCACAGAGGGGTAGTCTCTGCTGCTTCTCTGCTGCTTCTCTGCTGCTTCTCTGCT	480	1561	TCTGGGGCTTAGAGGCTGGGATGGCTCTGGGCGCCACCGCCAGAGGGGCGAAGCGCAGGCC	1620
QY	481	CATCTGCACAGAGGGGTAGTCTCTGCTGCTTCTCTGCTGCTTCTCTGCTGCTTCTCTGCT	540	1561	TCTGGGGCTTAGAGGCTGGGATGGCTCTGGGCGCCACCGCCAGAGGGGCGAAGCGCAGGCC	1620
Db	481	CATCTGCACAGAGGGGTAGTCTCTGCTGCTTCTCTGCTGCTTCTCTGCTGCTTCTCTGCT	540	1621	GGGCTGGAGGCGCGCGGCTCGGGTGGGGGGTCAAGTGGGACGCTGCTTCTCTCTCTCTCT	1680
QY	541	CATCTGCACAGAGGGGTAGTCTCTGCTGCTTCTCTGCTGCTTCTCTGCTGCTTCTCTGCT	600	1621	GGGCTGGAGGCGCGCGGCTCGGGTGGGGGGTCAAGTGGGACGCTGCTTCTCTCTCTCTCT	1680
Db	541	CATCTGCACAGAGGGGTAGTCTCTGCTGCTTCTCTGCTGCTTCTCTGCTGCTTCTCTGCT	600	1681	CTGCTGGCGCATCTCTTGGACCCAGATCTGTTCTGTTTCTGTTCTGTTCTGTTCTGTTCTG	1740
QY	601	GGAGGGCTGGCAGCTGTGTATCAAGGGCGCGGGGCTCCCATGCTGGGCTGTACAT	660	1681	CTGCTGGCGCATCTCTTGGACCCAGATCTGTTCTGTTTCTGTTCTGTTCTGTTCTGTTCTG	1740
Db	601	GGAGGGCTGGCAGCTGTGTATCAAGGGCGCGGGGCTCCCATGCTGGGCTGTACAT	660	1741	CTGCTGGCGCATCTCTTGGACCCAGATCTGTTCTGTTTCTGTTCTGTTCTGTTCTGTTCTG	1800
QY	661	GTGGACCTCTGTGGGACACAGAGAAATACAGCAAGTCACTGAGAGTTTCAGGTGC	720	1741	CTGCTGGCGCATCTCTTGGACCCAGATCTGTTCTGTTTCTGTTCTGTTCTGTTCTGTTCTG	1800
Db	661	GTGGACCTCTGTGGGACACAGAGAAATACAGCAAGTCACTGAGAGTTTCAGGTGC	720	1801	GCCCTTTGATGTCTTCTGATGAGGCTGCTTGGGCTGCTTGGGCTGCTTGGGCTGCTTGGG	1860
QY	721	AGAAACCCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	780	1801	GCCCTTTGATGTCTTCTGATGAGGCTGCTTGGGCTGCTTGGGCTGCTTGGGCTGCTTGGG	1860
Db	721	AGAAACCCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	780	1861	TTTGAAGAGTTTCCATGATCTGAAATCTGAAATCTGAAATCTGAAATCTGAAATCTGAAATCT	1920
QY	781	CTTCT	840	1861	TTTGAAGAGTTTCCATGATCTGAAATCTGAAATCTGAAATCTGAAATCTGAAATCTGAAATCT	1920
Db	781	CTTCT	840	1921	TCCAAAGTAGCTCCAAATGCAATTTTAAATGCAAAATCTGAAATCTGAAATCTGAAATCTG	1980
QY	841	ACGCGGAGAGAGAGTCTTCTCTCTAGAACCCAGATGAGGTCCGAGCCCTCAGAGC	900	1921	TCCAAAGTAGCTCCAAATGCAATTTTAAATGCAAAATCTGAAATCTGAAATCTGAAATCTG	1980
Db	841	ACGCGGAGAGAGAGTCTTCTCTCTAGAACCCAGATGAGGTCCGAGCCCTCAGAGC	900	1981	TAGTGGCTTTAAGCCCAAAAGCTCCCTAAGGCTCTCTGAGATGAAAGACGCGGGGGGAGC	2040
QY	901	GGAGGCCAGAGGGCTGAGCAGAGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	960	1981	TAGTGGCTTTAAGCCCAAAAGCTCCCTAAGGCTCTCTGAGATGAAAGACGCGGGGGGAGC	2040
Db	901	GGAGGCCAGAGGGCTGAGCAGAGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	960	2041	CCGAGCCAGGTGGAGACCCCGCAGGACGCGGGGGGCGGCTGAGCCGAGGCTTCGACACAG	2100
QY	961	CACCCCAAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1020	2041	CCGAGCCAGGTGGAGACCCCGCAGGACGCGGGGGGCGGCTGAGCCGAGGCTTCGACACAG	2100
Db	961	CACCCCAAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1020	2101	CCGCGCGCTTGGGGTCCGCGCGGAGCCAGGGTCCAGAGGGGCGGCTTGTGTCTCTCGGG	2160
QY	1021	GTCCCCCAACCTTGTTCATATGCGCGAGCCCATAGCCCTTCAGAGGAGAGAG	1080	2101	CCGCGCGCTTGGGGTCCGCGCGGAGCCAGGGTCCAGAGGGGCGGCTTGTGTCTCTCGGG	2160
Db	1021	GTCCCCCAACCTTGTTCATATGCGCGAGCCCATAGCCCTTCAGAGGAGAGAG	1080	2161	TTAAATAGGTTCCCGTCCG 2180	
QY	1081	GACAGGAGAGCCAGCTGAGTCCGACCTTGGGTGGCGGGGCTTGGGTCTCTCTCTCTCTCT	1140	2161	TTAAATAGGTTCCCGTCCG 2180	
Db	1081	GACAGGAGAGCCAGCTGAGTCCGACCTTGGGTGGCGGGGCTTGGGTCTCTCTCTCTCTCT	1140			
QY	1141	CACCCGAGGGCAGACACCGGCTTGTGTGGCAGGCTGGGCTCTGTGTCAACCCATCC	1200			

US-09-210-474-2
; Sequence 2, Application US/09210474
; Patent No. 6072034
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSES: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,474
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,559
; FILING DATE: 22-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2180 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-210-474-2

Query Match 100.0%; Score 2180; DB 3; Length 2180;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATTCCTGCTCTTTAGCGTGAACGCGGGTGGCGTCCCGTGAATAATAAATTCAC 60
DB 1 ATTCCTGCTCTTTAGCGTGAACGCGGGTGGCGTCCCGTGAATAATAAATTCAC 60
QY 61 CGTCACGCTGTGTGAACGCGGGTGGTCCCGAACTTGGAGGCTCCCGTAAACCCAG 120
DB 61 CGTCACGCTGTGTGAACGCGGGTGGTCCCGAACTTGGAGGCTCCCGTAAACCCAG 120
QY 121 CTCCTCTCTCATCTGGAGGTCGCGCGGGTCCCGAACTTGGAGGCTCCCGTAAACCCAG 180
DB 121 CTCCTCTCTCATCTGGAGGTCGCGCGGGTCCCGAACTTGGAGGCTCCCGTAAACCCAG 180
QY 181 CCTCTCGTGTCTTTTATTTTCTGGGGTCCGGGCGGAGAGCTGCATCCCAAGAGA 240
DB 181 CCTCTCGTGTCTTTTATTTTCTGGGGTCCGGGCGGAGAGCTGCATCCCAAGAGA 240
QY 241 GCGGCTCCAGGACGCGGAGTGTTCAGAGCCAGTGAAGAGCCAGAGGAGGAGGAG 300
DB 241 GCGGCTCCAGGACGCGGAGTGTTCAGAGCCAGTGAAGAGCCAGAGGAGGAGGAG 300
QY 301 AGTCCACACGACCATGCGACCTGCGCCCTGGCATTTCCCTGCGCACGTTTCCAGGCCCT 360
DB 301 AGTCCACACGACCATGCGACCTGCGCCCTGGCATTTCCCTGCGCACGTTTCCAGGCCCT 360
QY 361 TGGGACCCCTCTGTTTTTGGGCTGCTCTTTAGTGTCTCAGAAATGAAGGCTGGGACAGCCC 420

DB 361 TGGGACCCCTCTGTTTTTGGGCTGCTCTTGAAGTCTCAGAAATGAAGGCTGGGACAGCCC 420
QY 421 CATCTGCAACAGAGGGGTAGTCTCTGTGTCTTTGGGCGGAGAACCCGCTCATGTCTCTGCAA 480
DB 421 CATCTGCAACAGAGGGGTAGTCTCTGTGTCTTTGGGCGGAGAACCCGCTCATGTCTCTGCAA 480
QY 481 CATCTCAAACGCTCTCTCCCATGTCAACATCAAGCTGCGTCCCGGAGGAGAGCGC 540
DB 481 CATCTCAAACGCTCTCTCCCATGTCAACATCAAGCTGCGTCCCGGAGGAGAGCGC 540
QY 541 CATCTTCAATGAGGTGGCTCCAGGCTACTTCTCCCGGAGGAGGCTGGCAGCTCCAGGTTCA 600
DB 541 CATCTTCAATGAGGTGGCTCCAGGCTACTTCTCCCGGAGGAGGCTGGCAGCTCCAGGTTCA 600
QY 601 GGGAGGCGTGGCAGACAGCTGGTGATCAAAAGGCGCCCGGACTCCCATGTCTGGGCTGTACAT 660
DB 601 GGGAGGCGTGGCAGACAGCTGGTGATCAAAAGGCGCCCGGACTCCCATGTCTGGGCTGTACAT 660
QY 661 GTGGCACCTCTGTGGGACACAGAGAAATAACAGACAAGTCACTGGAGGTTTTCAGGTGC 720
DB 661 GTGGCACCTCTGTGGGACACAGAGAAATAACAGACAAGTCACTGGAGGTTTTCAGGTGC 720
QY 721 AGAACCCAGTCCGCGCCCTGACACTGGGTTCTGGGCTGTGGCAGGCGGTGCTACTGCTGT 780
DB 721 AGAACCCAGTCCGCGCCCTGACACTGGGTTCTGGGCTGTGGCAGGCGGTGCTACTGCTGT 780
QY 781 CTTTCATCTCTTGTGTGGTCTGTGGTCTGTGGTCTGTGGTCTGTGGTCTGTGGTCTGTGGT 840
DB 781 CTTTCATCTCTTGTGTGGTCTGTGGTCTGTGGTCTGTGGTCTGTGGTCTGTGGTCTGTGGT 840
QY 841 AGCGCGGAGAGAAAGTTCTTCTCTAGAACCCAGATGAAGGTCGACGCTTCAGAGC 900
DB 841 AGCGCGGAGAGAAAGTTCTTCTCTAGAACCCAGATGAAGGTCGACGCTTCAGAGC 900
QY 901 GGGAGCCAGACAGGCGCTGAGCAGAGCTCCGCTGAACTGTGGACCCAGACTCCGAGCC 960
DB 901 GGGAGCCAGACAGGCGCTGAGCAGAGCTCCGCTGAACTGTGGACCCAGACTCCGAGCC 960
QY 961 CACCCAAAGGCGCTGGCACTGGTCTCAAAACCCCTCACACTTGGAGCCCTGGAGCTGT 1020
DB 961 CACCCAAAGGCGCTGGCACTGGTCTCAAAACCCCTCACACTTGGAGCCCTGGAGCTGT 1020
QY 1021 GTCCCCCCTTGTTCATATGCGCAGACCCATAGCCGCTTGAAGGTCGACGAGAG 1080
DB 1021 GTCCCCCCTTGTTCATATGCGCAGACCCATAGCCGCTTGAAGGTCGACGAGAG 1080
QY 1081 GACACAGAGAGCAGCCCTGAGTCCGACCTTGGGTCGCGGGGCTTGGGCTCTCTGCTCC 1140
DB 1081 GACACAGAGAGCAGCCCTGAGTCCGACCTTGGGTCGCGGGGCTTGGGCTCTCTGCTCC 1140
QY 1141 CACCCGAGGCGCAGACACAGCGCTTGTGGCAGGCTGGGCTCTGTGTCAACCACTCC 1200
DB 1141 CACCCGAGGCGCAGACACAGCGCTTGTGGCAGGCTGGGCTCTGTGTCAACCACTCC 1200
QY 1201 TGGGTCGCTGACAGACCTTCCCTCCACCCCGCAGGCTTTCAGAGCTGTGCTCTCTCAGT 1260
DB 1201 TGGGTCGCTGACAGACCTTCCCTCCACCCCGCAGGCTTTCAGAGCTGTGCTCTCTCAGT 1260
QY 1261 TTCCAAATAGAAACCACTCTCCGACGACCCGACTTACAGAGAGCATGCCCTCC 1320
DB 1261 TTCCAAATAGAAACCACTCTCCGACGACCCGACTTACAGAGAGCATGCCCTCC 1320
QY 1321 CTCTGCCCTCATCAAAACCCAGACCTCCCTTCTGTGCGCCAGGCTGGTGGG 1380
DB 1321 CTCTGCCCTCATCAAAACCCAGACCTCCCTTCTGTGCGCCAGGCTGGTGGG 1380
QY 1381 CCCCAGGTCGTGGGTCGCTCTCTCCACTCCAGGCGCTCCGCGCCCAAGTGAAGGCGCC 1440
DB 1381 CCCCAGGTCGTGGGTCGCTCTCTCCACTCCAGGCGCTCCGCGCCCAAGTGAAGGCGCC 1440
QY 1441 CTGCGGAGCCTCAGACACACTGAGTTCAGGCTGGGCGGCTTGGCAGATACCTGTC 1500

1441 CTGCGGAGCCTCAGACACACTGGAGTTACGGGCTGGGGGGCCCTTGGCAATACCTGTC 1500
1501 CTTTGGCTATAGCAGGCTTTGGGGCCCTTCCGCGCAGCCCGGGGGCCGAGTAGGG 1560
1501 CTTTGGCTATAGCAGGCTTTGGGGCCCTTCCGCGCAGCCCGGGGGCCGAGTAGGG 1560
1561 TCTGGGGCTTAGAGGCTGGGATGCTCTGGGCCACCGCCAGGGGGCAAGCCAGGCC 1620
1561 TCTGGGGCTTAGAGGCTGGGATGCTCTGGGCCACCGCCAGGGGGCAAGCCAGGCC 1620
1621 GGGCTGGAGGGCGGGGGGGCTCGGGCTGGGGGGTCAAGTGAACCTGCCTCCGGGG 1680
1621 GGGCTGGAGGGCGGGGGGGCTCGGGCTGGGGGGTCAAGTGAACCTGCCTCCGGGG 1680
1681 CTGGTGGCGCATCCCTCAGTCCCTCGGCCACCGGGGGTGCCTCCTCGTGCACCGCA 1740
1681 CTGGTGGCGCATCCCTCAGTCCCTCGGCCACCGGGGGTGCCTCCTCGTGCACCGCA 1740
1741 CTTGGGAGCCTTTTGGACCCAGATCTGTTATGCTTTTGTCTGCTGCTGCGGGGG 1800
1741 CTTGGGAGCCTTTTGGACCCAGATCTGTTATGCTTTTGTCTGCTGCTGCGGGGG 1800
1801 GCCTTTGATGCTTTCATCTGTATGGGTGGGAAATACCGGGGAATCCCGCTTCAGTTC 1860
1801 GCCTTTGATGCTTTCATCTGTATGGGTGGGAAATACCGGGGAATCCCGCTTCAGTTC 1860
1861 TTTGAAAAAGTTCCATGACTCGAATATCGAATGAAGAAACAAACCGACTCACAAACC 1920
1861 TTTGAAAAAGTTCCATGACTCGAATATCGAATGAAGAAACAAACCGACTCACAAACC 1920
1921 TCCAAGTAGCTCCAAATCAATTTTAAATGGAACAAATAATCTGAAGAAACGCTTT 1980
1921 TCCAAGTAGCTCCAAATCAATTTTAAATGGAACAAATAATCTGAAGAAACGCTTT 1980
1981 TAGTGGCTTTAAGCCCAAAACGCTCCTAAGGCGTCTCGAGATGAAGACGGGGGGAGC 2040
1981 TAGTGGCTTTAAGCCCAAAACGCTCCTAAGGCGTCTCGAGATGAAGACGGGGGGAGC 2040
2041 CCCAGCAGGTGAGACCCCGCAGGACGGCGGCGCCCGGTGACCGAGGCTTCGCACAG 2100
2041 CCCAGCAGGTGAGACCCCGCAGGACGGCGGCGCCCGGTGACCGAGGCTTCGCACAG 2100
2101 CCGGCGCCCTGAGGGTGGGGCGGAGCCAGGCTTCAAGAGGGCGGCTTTGTCTCGGG 2160
2101 CCGGCGCCCTGAGGGTGGGGCGGAGCCAGGCTTCAAGAGGGCGGCTTTGTCTCGGG 2160
2161 TTAATAAAGGTTCCGTCG 2180
2161 TTAATAAAGGTTCCGTCG 2180

RESULT 3
US-09-539-774-2
; Sequence 2, Application US/09539774
; Patent No. 6350615
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P. C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/539,774
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/210,474
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2180 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-539-774-2
Query Match 100.0%; Score 2180; DB 3; Length 2180;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATTCCTGCTTCTTTAGCGTGAACCGGGTGGGCTGCTCCCGTGAATAATAAATTCAC 60
DB 1 ATTCCTGCTTCTTTAGCGTGAACCGGGTGGGCTGCTCCCGTGAATAATAAATTCAC 60
QY 61 CGTCACGCTTCTTGTGAACCGGGTGGTTCGCAAACTTTGGAGGCTTCCCGTAAACCCAG 120
DB 61 CGTCACGCTTCTTGTGAACCGGGTGGTTCGCAAACTTTGGAGGCTTCCCGTAAACCCAG 120
QY 121 CTCCTTCTCATCTGGGAGGTGGTCCCGCGCGGGTCCCGCGCTCTCTCCCTGCGCCCTC 180
DB 121 CTCCTTCTCATCTGGGAGGTGGTCCCGCGCGGGTCCCGCGCTCTCTCCCTGCGCCCTC 180
QY 181 CTCCTCGTGTCTTTCAATTTCTTGGGGCTCCGGGGCGCGAGAGCTGCATCCAGAGGA 240
DB 181 CCTCTGCTGTCTTTCAATTTCTTGGGGCTCCGGGGCGCGAGAGCTGCATCCAGAGGA 240
QY 241 GCGGCTCAGAGCGGACCCCGGAGTGTTCAGAGCCAGTCAAGAGCACCGAGGCGGCCA 300
DB 241 GCGGCTCAGAGCGGACCCCGGAGTGTTCAGAGCCAGTCAAGAGCACCGAGGCGGCCA 300
QY 301 AGTCCACACGACATGACAGCTGCGCCCTCGCATTTCCCTGGCCACGTTTCCAGGCCCT 360
DB 301 AGTCCACACGACATGACAGCTGCGCCCTCGCATTTCCCTGGCCACGTTTCCAGGCCCT 360
QY 361 TGGGACCCCTCTTGTGTTTGGTGTCTCTGAGTGTCTAGATGAAGGCTGGGACAGCCC 420
DB 361 TGGGACCCCTCTTGTGTTTGGTGTCTCTGAGTGTCTAGATGAAGGCTGGGACAGCCC 420
QY 421 CATCTGCACAGAGGGGTAGTCTCTGCTTTGGGCGGAGAACACCGTCTCTCTGCAA 480
DB 421 CATCTGCACAGAGGGGTAGTCTCTGCTTTGGGCGGAGAACACCGTCTCTCTGCAA 480
QY 481 CATCTCCAAACGCTTCTCCCATGTCAACATCAAGCTGCGTCCCAAGAGGAGAGCC 540
DB 481 CATCTCCAAACGCTTCTCCCATGTCAACATCAAGCTGCGTCCCAAGAGGAGAGCC 540
QY 541 CATCTTCAATGAGGTGCTCCAGGCTACTTCTCCCGGGACGGCTGGAGTTCA 600
DB 541 CATCTTCAATGAGGTGCTCCAGGCTACTTCTCCCGGGACGGCTGGAGTTCA 600
QY 601 GGGAGGCTGGCACAGCTGTGATCAAGCGCGCCGGGACTCCCATGTCTGGGTGTACAT 660
DB 601 GGGAGGCTGGCACAGCTGTGATCAAGCGCGCCGGGACTCCCATGTCTGGGTGTACAT 660
QY 661 GTGGCACCTCGTGGGACACAGAGAAATACAGAGTCAAGCTGAGGTTTCAAGTGC 720
DB 661 GTGGCACCTCGTGGGACACAGAGAAATACAGAGTCAAGCTGAGGTTTCAAGTGC 720

Db 661 GTGGCACCTCTGTGGGACACGAGAAATAACAGCAAGTCA CGCTGGAGGTTTCAGGTGC 720
Qy 721 AGAACCCAGTCCGCCCTGACACTGGGTTCTGGCCTGTGCGACAGCGGTGGTCACTGCTGT 780
Db 721 AGAACCCAGTCCGCCCTGACACTGGGTTCTGGCCTGTGCGACAGCGGTGGTCACTGCTGT 780
Qy 781 CTTTCATCTCTTGGTGGCTCTGCTGATGTTGGCTGGTACAGGTGCGCTGTTCAGCA 840
Db 781 CTTTCATCTCTTGGTGGCTCTGCTGATGTTGGCTGGTACAGGTGCGCTGTTCAGCA 840
Qy 841 AGCGGGGAGAGAGAGTCTCTCTCTAGAACCCAGATGAAGTTCGACAGCCCTCAGAGC 900
Db 841 AGCGGGGAGAGAGAGTCTCTCTCTAGAACCCAGATGAAGTTCGACAGCCCTCAGAGC 900
Qy 901 GGGAGCCAGCAGGGGCTGAGCAGAGCCTCCGCTGAACCTGTGGACCCAGACTCCGAGCC 960
Db 901 GGGAGCCAGCAGGGGCTGAGCAGAGCCTCCGCTGAACCTGTGGACCCAGACTCCGAGCC 960
Qy 961 CACCCAGAGCGCTGGACTGTTTCAACCTCACCCTTGGAGCCCTGGAGTGTCT 1020
Db 961 CACCCAGAGCGCTGGACTGTTTCAACCTCACCCTTGGAGCCCTGGAGTGTCT 1020
Qy 1021 GTCCCCCCCAACCTTGTTCATATGCCGAGACCCATAGCCGCTGCAAGGACAGAG 1080
Db 1021 GTCCCCCCCAACCTTGTTCATATGCCGAGACCCATAGCCGCTGCAAGGACAGAG 1080
Qy 1081 GACAGAGAGCCAGCCCTGAGTGCAGACTTGGGTGGCGGGGCTCTCGTCC 1140
Db 1081 GACAGAGAGCCAGCCCTGAGTGCAGACTTGGGTGGCGGGGCTCTCGTCC 1140
Qy 1141 CACCCGAGGACAGACACCGGCTTGTGGAGGCTGGGCTCTGTGTCAACCACTCC 1200
Db 1141 CACCCGAGGACAGACACCGGCTTGTGGAGGCTGGGCTCTGTGTCAACCACTCC 1200
Qy 1201 TGGGTCGTGACAGACCTTCCCTCCACCCCGAGTCTTCAAGCTCTGCTCTCACT 1260
Db 1201 TGGGTCGTGACAGACCTTCCCTCCACCCCGAGTCTTCAAGCTCTGCTCTCACT 1260
Qy 1261 TTCCAAATGGAACCACTCACTCCGACGACCCGACTTACGAGCAGCATGCCCTCC 1320
Db 1261 TTCCAAATGGAACCACTCACTCCGACGACCCGACTTACGAGCAGCATGCCCTCC 1320
Qy 1321 CTTGCTCTCATCAACCCAGACACCGGACTTCCCTTCTGCAACCCAGGCTGTTCGG 1380
Db 1321 CTTGCTCTCATCAACCCAGACACCGGACTTCCCTTCTGCAACCCAGGCTGTTCGG 1380
Qy 1381 CCCCAGGTGTGGGTCGGCTCTCCACTCCAGGCTCCGCGCCCAAGTGAAGGGGCC 1440
Db 1381 CCCCAGGTGTGGGTCGGCTCTCCACTCCAGGCTCCGCGCCCAAGTGAAGGGGCC 1440
Qy 1441 CTGCGGAGCCTCAGACACACTGGAGTTCAAGGCTGGGGGCTTGGACATACCTGTC 1500
Db 1441 CTGCGGAGCCTCAGACACACTGGAGTTCAAGGCTGGGGGCTTGGACATACCTGTC 1500
Qy 1501 CTTGGCTATAGAGGCTTGGGGCCCTTCCGGGACAGCCCGGGGGCCGAGGTAGGG 1560
Db 1501 CTTGGCTATAGAGGCTTGGGGCCCTTCCGGGACAGCCCGGGGGCCGAGGTAGGG 1560
Qy 1561 TCTGGGGCTTAGAGGCTGGGATGCTCTGCGCCACCGCCAGGGGGCAGCGAGGCC 1620
Db 1561 TCTGGGGCTTAGAGGCTGGGATGCTCTGCGCCACCGCCAGGGGGCAGCGAGGCC 1620
Qy 1621 GGGCTGGAGGGGGGGGGCTGGGCTGGGGCTGGGGCTGAGTGAAGTCTCGGG 1680
Db 1621 GGGCTGGAGGGGGGGGGCTGGGCTGGGGCTGGGGCTGAGTGAAGTCTCGGG 1680
Qy 1681 CTGGTCGCGCATCTCTAGTCTCTGGCCACCGGGGGTGTCTCTGTCGCCACCGCA 1740
Db 1681 CTGGTCGCGCATCTCTAGTCTCTGGCCACCGGGGGTGTCTCTGTCGCCACCGCA 1740
Qy 1741 CTTGCGGAGCCTCTTTGGACCCAGATCTGTTATGCTTTTGTCTGCTGCTGGGG 1800
Db 1741 CTTGCGGAGCCTCTTTGGACCCAGATCTGTTATGCTTTTGTCTGCTGCTGGGG 1800

Qy 1801 GCCCTTTGATGTCTTTCATCTGTATGGGTGGAAAAATCACCGGAATCCCTTTCAGTTTC 1860
Db 1801 GCCCTTTGATGTCTTTCATCTGTATGGGTGGAAAAATCACCGGAATCCCTTTCAGTTTC 1860
Qy 1861 TTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAAAACAACCGACTCACAAC 1920
Db 1861 TTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAAAACAACCGACTCACAAC 1920
Qy 1921 TCCAAGTAGTCCAAATGCAATTTTAAATGAAAAAACAATACTGAAAGAAACGCTTT 1980
Db 1921 TCCAAGTAGTCCAAATGCAATTTTAAATGAAAAAACAATACTGAAAGAAACGCTTT 1980
Qy 1981 TAGTGGCTTTAAGCCCCAAACGTCCTTAAGCGCTCTCGAGATGAAGACGGGGGGAGC 2040
Db 1981 TAGTGGCTTTAAGCCCCAAACGTCCTTAAGCGCTCTCGAGATGAAGACGGGGGGAGC 2040
Qy 2041 CCCAGCAGGTGAGACCCCGCAGACCGCGGGCGCGCGGTGACCGAGGCTTCGCACAG 2100
Db 2041 CCCAGCAGGTGAGACCCCGCAGACCGCGGGCGCGCGGTGACCGAGGCTTCGCACAG 2100
Qy 2101 CCGGCGCCCTGAGGCTCGGGCCGAGCCAGGCTTCAAGAGGGCGCGCTTGTGTCTCGGG 2160
Db 2101 CCGGCGCCCTGAGGCTCGGGCCGAGGCTTCAAGAGGGCGCGCTTGTGTCTCGGG 2160
Qy 2161 TTAATAAAGGTTCCGTCGG 2180
Db 2161 TTAATAAAGGTTCCGTCGG 2180

RESULT 4

US-09-997-165-3
; Sequence 3, Application US/09997165
; Patent No. 6762030
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; PRIORITY FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(865)
US-09-997-165-3

Query Match 89.5%; Score 1951; DB 4; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
Qy 196 ATTTTCTTGGGGCTCCGGGGCGCGAGAACTGCAATCCAGAGGAGCGGTCCAGAGCG 255
Db 1 ATTTTCTTGGGGCTCCGGGGCGCGAGAACTGCAATCCAGAGGAGCGGTCCAGAGCG 60
Qy 256 GACCCGGGAGTGTTCAGAGCCAGTGAAGGACAGGGGGCCCAAGTCCACACCAAT 315
Db 61 GACCCGGGAGTGTTCAGAGCCAGTGAAGGACAGGGGGCCCAAGTCCACACCAAT 120
Qy 316 GCAGACCTGCCCCCTGGCAATTCCTTGGCCACGCTTTCCAGGGCCCTTGGAGCCCTCTGTT 375
Db 121 GCAGACCTGCCCCCTGGCAATTCCTTGGCCACGCTTTCCAGGGCCCTTGGAGCCCTCTGTT 180
Qy 376 TTTGGCTGCTCTCTTGGAGTGTCTCAGAAATGAAGGCTGGGACAGCCCCCATCTGCACAGAGG 435

181	TTTGCTGCTTCTTGTAGTCTCAGATGAGGCTGGACAGCCCATCTGCACAGAGG	240	1260	ACACACTGAGTTTCTGAGGCTGGGGGGGCTTTGGCACATACCTGTCCCTTGGCTATGAGCA	1319
436	GGTAGTCTCTGTGTTTGGGGGAGAACACCGTCTATGTCCTGCAACATCTCCACGCTT	495	1516	GGCTTTGGGGGGCTTCCGGCGACCCCGGGGGCCGAGGTAGGCTCTGGGGGCTTAGAG	1575
241	GGTAGTCTCTGTGTTTGGGGGAGAACACCGTCTATGTCCTGCAACATCTCCACGCTT	300	1320	GGCTTTGGGGGGCTTCCGGCGACCCCGGGGGCCGAGGTAGGCTCTGGGGGCTTAGAG	1379
496	CTCCATGTCACATCAAGTCTGCTGCGCCACCGGCGAGGAGCGCCATCTTCAATGAGGT	555	1576	GCTGGGATGCTCTCTGGGCCCAACCGCCAGGGGGCAAGCGAGGCGGCTGGAGCGGC	1635
301	CTCCATGTCACATCAAGTCTGCTGCGCCACCGGCGAGGAGCGCCATCTTCAATGAGGT	360	1380	GCTGGGATGCTCTCTGGGCCCAACCGCCAGGGGGCAAGCGAGGCGGCTGGAGCGGC	1439
556	GGCTCCAGGCTACTTCTCCCGGGACGGCTGGAGCTCCAGGTTGAGGAGCGCTGGACA	615	1636	GGCGGGCTCTGGGGGCTGGGGGGTCAAGTGGAGCGCTCCGGGGCTGGTCGCGCATCCC	1695
361	GGCTCCAGGCTACTTCTCCCGGGACGGCTGGAGCTCCAGGTTGAGGAGCGCTGGACA	420	1440	GGCGGGCTCTGGGGGCTGGGGGGTCAAGTGGAGCGCTCCGGGGCTGGTCGCGCATCCC	1499
616	GCTGGTGTATCAAGGGGCGCGGACTCCCATGCTGGGCTGTACATGTGGACCTCGTGG	675	1696	TCAGTCCCTCGGCGACCCGGGGGCTCGCTCCCTCGTGGCGCCACCGACCTGCCGAGCTCTT	1755
421	GCTGGTGTATCAAGGGGCGCGGACTCCCATGCTGGGCTGTACATGTGGACCTCGTGG	480	1500	TCAGTCCCTCGGCGACCCGGGGGCTCGCTCCCTCGTGGCGCCACCGACCTGCCGAGCTCTT	1559
676	ACACACAGAGAAATAACAGACAAAGTCAACGCTGGAGGTTTCAAGTGCAGAACCCAGTCCGC	735	1756	TGGACCCAGATCTGTTTCATGCTTTTGTCTTGTCTGTCACCTGGGGGGGCTTTGATGTCTT	1815
481	ACACACAGAGAAATAACAGACAAAGTCAACGCTGGAGGTTTCAAGTGCAGAACCCAGTCCGC	540	1560	TGGACCCAGATCTGTTTCATGCTTTTGTCTTGTCTGTCACCTGGGGGGGCTTTGATGTCTT	1619
736	CCCTGACACTGGGTTCTGGCTGTGCGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT	795	1816	CATCTGTATGGGTGGGAAATCACCGGAAATCCCTCTTCAAGTCTTTGAAAAAGTTCCA	1875
541	CCCTGACACTGGGTTCTGGCTGTGCGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT	600	1620	CATCTGTATGGGTGGGAAATCACCGGAAATCCCTCTTCAAGTCTTTGAAAAAGTTCCA	1679
796	CGCTCTGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	855	1876	TGACTCGAATATCTGAATGAAGAAACAAACCGACTCAACACCTCCAGTAGCTCCAA	1935
601	CGCTCTGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	660	1680	TGACTCGAATATCTGAATGAAGAAACAAACCGACTCAACACCTCCAGTAGCTCCAA	1739
856	GTTCTTCTCTAGAACCCAGATGAGTTCAGGCTTCAGAGCGGGAGCCAGCAGGG	915	1936	ATGCAATTTTAAATGGAACCAAAATCTGAAGAAACGCTCTTTAGTGGCTTTAAGCC	1995
661	GTTCTTCTCTAGAACCCAGATGAGTTCAGGCTTCAGAGCGGGAGCCAGCAGGG	720	1740	ATGCAATTTTAAATGGAACCAAAATCTGAAGAAACGCTCTTTAGTGGCTTTAAGCC	1799
916	CCTGACAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	975	1996	CCAAAACTCCTTAAGCGCTCTCGAGATGAAGACCGGGGGGAG - CCCCCAGCAGGTGGA	2054
721	CCTGACAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	780	1800	CCAAAACTCCTTAAGCGCTCTCGAGATGAAGACCGGGGGGAG - CCCCCAGCAGGTGGA	1859
976	GGCACTGTGTTTCAAAACCTTACACATTTGAGGCTTGGAGCTGCTGCTGCTGCTGCTGCTGCT	1035	2055	GACCCCGCAGACCGCGGGCGCGCTGACCGAGGCTCGCACAGCGCGCGCTCGAG	2114
781	GGCACTGTGTTTCAAAACCTTACACATTTGAGGCTTGGAGCTGCTGCTGCTGCTGCTGCTGCT	839	1860	GACCCCGCAGACCGCGGGCGCGCTGACCGAGGCTCGCACAGCGCGCGCTCGAG	1919
1036	TGTTTCCATATGCGCGACCCATAGCGGCTTCAAGGCGAGGAGGACACAGAGAGCCA	1095	2115	GTCGCGGC - GAGCCAGGCTCAAGAGGGGCGCGTTTGTGTCCTCGGTTAAAAAAGGTT	2173
840	TGTTTCCATATGCGCGACCCATAGCGGCTTCAAGGCGAGGAGGACACAGAGAGCCA	899	1920	GTCGCGGC - GAGCCAGGCTCAAGAGGGGCGCGTTTGTGTCCTCGGTTAAAAAAGGTT	1979
1096	GCCCTGAGTGGCGACTTGGGTGGCGGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1155	2174	CCGTCCG 2180	
900	GCCCTGAGTGGCGACTTGGGTGGCGGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	959	1980	CCGTCCG 1986	
1156	GACACCGGCTTGTGGAGGCTGGGCTCTGTGTACCCACTCTGCTGCTGCTGCTGCTGCTGCTGCT	1215			
960	GACACCGGCTTGTGGAGGCTGGGCTCTGTGTACCCACTCTGCTGCTGCTGCTGCTGCTGCTGCT	1019			
1216	CCTTCCCTTCCACCCCGAGGTTTCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1275			
1020	CCTTCCCTTCCACCCCGAGGTTTCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1079			
1276	ACCTACCTTCCGACGACCCGACTTACAGGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1335			
1080	ACCTACCTTCCGACGACCCGACTTACAGGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1139			
1336	ACCCAGACCGGACTCTCTTCTGCGACCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1395			
1140	ACCCAGACCGGACTCTCTTCTGCGACCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1199			
1396	CGCTCTCTCACTCCAGGCTTCCGCGCCCAAGTGAAGGGGCGCTGCGGAGGCTCAG	1455			
1200	CGCTCTCTCACTCCAGGCTTCCGCGCCCAAGTGAAGGGGCGCTGCGGAGGCTCAG	1259			
1456	ACACACTGAGGTTTCTGAGGCTGGGGGGCTTGGACATACCTGCTCCCTGGCTATGAGCA	1515			

RESULT 5
US-09-949-016-426
; Sequence 426, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 426
; LENGTH: 2000
; TYPE: DNA

ORGANISM: Human
US-09-949-016-426

Query Match 89.5%; Score 1951; DB 4; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

Qy	196	ATTTCCTGGGGCTCCGGGCGCGAGAGCTGCAATCCAGAGGAGCGCTCCAGAGCG	255
Db	1	ATTTCCTGGGGCTCCGGGCGCGAGAGCTGCAATCCAGAGGAGCGCTCCAGAGCG	60
Qy	256	GACCCGGAGTGTTCAGAGCCAGTGAAGGACAGGGGCCCAAGTCCACAGCCAT	315
Db	61	GACCCGGAGTGTTCAGAGCCAGTGAAGGACAGGGGCCCAAGTCCACAGCCAT	120
Qy	316	GCAGACCTGCCCTCCCTGGCCATCCCTGGCCACCGTTCACAGGCGCTTGGACCTCTCTGT	375
Db	121	GCAGACCTGCCCTCCCTGGCCATCCCTGGCCACCGTTCACAGGCGCTTGGACCTCTCTGT	180
Qy	376	TTTGGCTGCTCCTTTGAGTGTCTAGAATGAAGCTGGGACAGCCCATCTGACAGAGGG	435
Db	181	TTTGGCTGCTCCTTTGAGTGTCTAGAATGAAGCTGGGACAGCCCATCTGACAGAGGG	240
Qy	436	GGTAGTCTCTGTCTTTGGGGCGAGAACACCGTCAATGCTCTGCAACATCTCCAGCCCTT	495
Db	241	GGTAGTCTCTGTCTTTGGGGCGAGAACACCGTCAATGCTCTGCAACATCTCCAGCCCTT	300
Qy	496	CTCCCATGTCAACATCAAGCTGCGTCCACCGGGCAGGAGCGGCATCTTCAATGAGGT	555
Db	301	CTCCCATGTCAACATCAAGCTGCGTCCACCGGGCAGGAGCGGCATCTTCAATGAGGT	360
Qy	556	GGCTCCAGGCTACTTCTCCGGGACGGCTGGAGCTCCAGGTTCCAGGAGCGTGGCACA	615
Db	361	GGCTCCAGGCTACTTCTCCGGGACGGCTGGAGCTCCAGGTTCCAGGAGCGTGGCACA	420
Qy	616	GCTGGTATCAAGCGCCCGGACTCCCATGCTGGGCTGTACATGTGGACCTCGTGGG	675
Db	421	GCTGGTATCAAGCGCCCGGACTCCCATGCTGGGCTGTACATGTGGACCTCGTGGG	480
Qy	676	ACACGAGAGAAATACAGACAAAGTCACTGGAGGTTTCAGGTGAGAACCCAGTCCGC	735
Db	481	ACACGAGAGAAATACAGACAAAGTCACTGGAGGTTTCAGGTGAGAACCCAGTCCGC	540
Qy	736	CCCTGACACTGGGTTCTGGCTGTGCCAGCGGTGGTCACTGCTGTTTCATCTCTGGT	795
Db	541	CCCTGACACTGGGTTCTGGCTGTGCCAGCGGTGGTCACTGCTGTTTCATCTCTGGT	600
Qy	796	CGCTTGCTCATGTTCCGCTGTGACAGTGGCGGTGTTCCAGCMAACCGCGGAGAGAA	855
Db	601	CGCTTGCTCATGTTCCGCTGTGACAGTGGCGGTGGTCACTGCTGTTTCATCTCTGGT	660
Qy	856	GTTCTTCTCTAGAACCCAGATGAAGTGCAGCCCTCAGAGCGGAGCCAGCAGGG	915
Db	661	GTTCTTCTCTAGAACCCAGATGAAGTGCAGCCCTCAGAGCGGAGCCAGCAGGG	720
Qy	916	CCTGAGCAGAGCTTCGCTGAACTGTGGACCCCGAGCTCCGAGGCCACCCCAAGCCGCT	975
Db	721	CCTGAGCAGAGCTTCGCTGAACTGTGGACCCCGAGCTCCGAGGCCACCCCAAGCCGCT	780
Qy	976	GGCACTGGTGTTCAAACCTCACCCTTGGAGCCCTGGAGCTGTGTCTCCCGCCCAACCCCT	1035
Db	781	GGCACTGGTGTTCAAACCTCACCCTTGGAGCCCTGGAGCTGTGTCTCCCGCCCAACCCCT	839
Qy	1036	TGTTTCCATATGCGCAGACCCATAGCGCTGCAAGGCTGAGAGGACACAGGAGGCA	1095
Db	840	TGTTTCCATATGCGCAGACCCATAGCGCTGCAAGGCTGAGAGGACACAGGAGGCA	899
Qy	1096	GCCCTGAGTCCGACCTTGGGTGGGGGCGCTGGGTCTCTGCTCCACCCCGAGGAGGACA	1155
Db	900	GCCCTGAGTCCGACCTTGGGTGGGGGCGCTGGGTCTCTGCTCCACCCCGAGGAGGACA	959
Qy	1156	GACACCGGCTTGTGGAGCGCTGGGCTCTGTGTACCCACTCTCTCTGGGTGGTGCAGAC	1215

RESULT 6
US-09-949-016-2936

; Sequence 2936, Application US/09949016									
; Patent No. 6812339									
; GENERAL INFORMATION:									
; APPLICANT: VENTER, J. Craig et al.									
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED									
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF									
; FILE REFERENCE: CL001307									
; CURRENT APPLICATION NUMBER: US/09/949,016									
; CURRENT FILING DATE: 2000-04-14									
; PRIOR APPLICATION NUMBER: 60/241,755									
; PRIOR FILING DATE: 2000-10-20									
; PRIOR APPLICATION NUMBER: 60/237,768									
; PRIOR FILING DATE: 2000-10-03									
; PRIOR APPLICATION NUMBER: 60/231,498									
; PRIOR FILING DATE: 2000-09-08									
; NUMBER OF SEQ ID NOS: 207012									
; SOFTWARE: FastSeq for Windows Version 4.0									
; SEQ ID NO 2936									
; LENGTH: 2001									
; TYPE: DNA									
; ORGANISM: Human									
US-09-949-016-2936									
Query Match 86.7%; Score 1889.6; DB 4; Length 2001;									
Best Local Similarity 99.1%; Pred. No. 0;									
Matches 1973; Conservative 0; Mismatches 9; Indels 8; Gaps 7;									
QY	196	ATTTTCCTCGGGCTCGGGGCGGAGAGAGCTGCATCCAGAGGAGCGCGTCCAGAGCG	255						
DB	1	ATTTTCCTCGGGCTCGGGGCGGAGAGAGCTGCATCCAGAGGAGCGCGTCCAGAGCG	60						
QY	256	GACCCGGGAGTGTTCAGAGCCAGTGACAGACAGAGGAGGAGGAGGAGGAGGAGGAGG	315						
DB	61	GACCCGGGAGTGTTCAGAGCCAGTGACAGACAGAGGAGGAGGAGGAGGAGGAGGAGG	120						
QY	316	GCAGACTGCTCCCTGCGCATTCCTGCGCAGCTTCCAGAGGAGGAGGAGGAGGAGGAGG	375						
DB	121	GCAGACTGCTCCCTGCGCATTCCTGCGCAGCTTCCAGAGGAGGAGGAGGAGGAGGAGG	180						
QY	376	TTTGGCTGCTCCTTTGAGTGTCTCAGAAATGAAGCTGGGAGGAGGAGGAGGAGGAGG	435						
DB	181	TTTGGCTGCTCCTTTGAGTGTCTCAGAAATGAAGCTGGGAGGAGGAGGAGGAGGAGG	240						
QY	436	GTTAGTCTCTGTGTTTGGGGGAGAGACCGCTCATGCTCTGCAACATCTCCAAAGCCTT	495						
DB	241	GTTAGTCTCTGTGTTTGGGGGAGAGACCGCTCATGCTCTGCAACATCTCCAAAGCCTT	300						
QY	496	CTCCCATGTCAACATCAAGCTGCGTGCCTCCAGCGGAGGAGGAGGAGGAGGAGGAGG	555						
DB	301	CTCCCATGTCAACATCAAGCTGCGTGCCTCCAGCGGAGGAGGAGGAGGAGGAGGAGG	360						
QY	556	GGCTCCAGGCTACTTCTCCGGGAGCGCTGCAGCTCCAGGTTTCAGGAGGAGGAGGAGG	615						
DB	361	GGCTCCAGGCTACTTCTCCGGGAGCGCTGCAGCTCCAGGTTTCAGGAGGAGGAGGAGG	420						
QY	616	GCTGGTATCAAGCGCGCGGAGCTCCATGCTGGGCTGTACATGTGGGAGGAGGAGGAGG	675						
DB	421	GCTGGTATCAAGCGCGCGGAGCTCCATGCTGGGCTGTACATGTGGGAGGAGGAGGAGG	480						
QY	676	ACACCAGAGAAATACAGACAAGTCAAGCTGGAGGTTTCAGGTTTCAGGAGGAGGAGG	735						
DB	481	ACACCAGAGAAATACAGACAAGTCAAGCTGGAGGTTTCAGGTTTCAGGAGGAGGAGG	540						
QY	736	CCCTGACACTGGGTTCTGGCTGTGCGCAGCGGTGCTGCTGCTTTCATCTCTTGGT	795						
DB	541	CCCTGACACTGGGTTCTGGCTGTGCGCAGCGGTGCTGCTGCTTTCATCTCTTGGT	600						
QY	796	CGCTCTGCTGATGTTTCGGCTGTGACAGGTGCGCTGTTTCCAGCAACCGCGGAGAGAA	855						
DB	601	CGCTCTGCTGATGTTTCGGCTGTGACAGGTGCGCTGTTTCCAGCAACCGCGGAGAGAA	660						
QY	856	GTTCTTCTCTCTAGAACCCCATGCAAGGTCGACGCTCTCAGAGGAGGAGGAGGAGG	915						
DB									

661 GTTCTTCTCTAGAAACCCAGATGAAGTGTGAGCCCTCAGAGCGGAGCCAGAGGG 720

916 CTGTAGCAGAGCTCGCTGAACTGTGGACCCAGACTCCGAGCCACACCCCAAGCGCT 975

721 CTGTAGCAGAGCTCGCTGAACTGTGGACCCAGACTCCGAGCCACACCCCAAGCGCT 780

976 GGCACCTGGTGTTCAAACCCCTCACCCTTGGAGCCCTGGAGCTGCTGTCCCCCACCCT 1035

781 GGCACCTGGTGTTCAAACCCCTCACCCTTGGAGCCCTGGAGCTGCTGTCCCCCACCCT 839

1036 TGTTCATATGCGCAGACCCCATAGCCGCTGCAAGGAGGAGGAGGAGGAGGAGGAGG 1095

840 TGTTCATATGCGCAGACCCCATAGCCGCTGCAAGGAGGAGGAGGAGGAGGAGGAGG 899

1096 GGCCTGAGTGCAGACCTTGGTGGCGGGGCTGGGCTCTCGTCCACCCGAGGAGGAGG 1155

900 GGCCTGAGTGCAGACCTTGGTGGCGGGGCTGGGCTCTCGTCCACCCGAGGAGGAGG 959

1156 GACACCGGCTTGTGTGGCAGGCTGGGCTTGTGTACCCACTCTCTGGGTGGTGCAGAC 1215

960 GACACCGGCTTGTGTGGCAGGCTGGGCTTGTGTACCCACTCTCTGGGTGGTGCAGAC 1019

1216 CTTTCCCTCCACCCCGAGGCTTCCAGCTCTGCTTCTCAGTTTCCAAATGGAAC 1275

1020 CTTTCCCTCCACCCCGAGGCTTCCAGCTCTGCTTCTCAGTTTCCAAATGGAAC 1079

1276 ACCTCACCTCGGAGCACCAGCTTACAGAGCAGCATGCCCCCTCTCTGCCCTCATCA 1335

1080 ACCTCACCTCGGAGCACCAGCTTACAGAGCAGCATGCCCCCTCTCTGCCCTCATCA 1139

1336 ACCCAGAGCCGAGCTCCCTTCTGCGCAGCCAGGCTGTCGGGCCCGCAGGTTGGGGT 1395

1140 ACCCAGAGCCGAGCTCCCTTCTGCGCAGCCAGGCTGTCGGGCCCGCAGGTTGGGGT 1199

1396 CCGCTCTCTCCTCCTCCAGGCTCCGCGCCCAAGTGAAGGGGCGCTTCCGCGAGCTCAG 1455

1200 CCGCTCTCTCCTCCTCCAGGCTCCGCGCCCAAGTGAAGGGGCGCTTCCGCGAGCTCAG 1259

1456 ACACACTCGAGTTCAGGGC--TGGGGGGGCTTGGGCAATACCTCTCTGGCTATGAG 1513

1260 ACACACTCGAGTTCAGGGCTTGGGGGGCTTGGGCAATACCTCTCTGGCTATGAG 1319

1514 CAGGCTTGGGGGGCTTCCGCGCAGCCCGGGGCGGAGGTAGGCTCTGGGGGCTTGG 1573

1320 CAGGCTTGGGGGGCTTCCGCGCAGCCCGGGGCGGAGGTAGGCTCTGGGGGCTTGG 1378

1574 AGGCTGGGATGGCTCCTGCGCCCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1633

1379 AGGCTGGGATGGCTCCTGCGCCCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1437

1634 GCGCGGCGGCTCGGGCTGGGGGCTCAGGTGAGCGCTG-CCTCCGGGGCTGGTCCGCGAT 1692

1438 GCGCGGCGGCTCGGGCTGGGGGCTCAGGTGAGCGCGCGCTTCCGGGGCTGGAGCGCAT 1497

1693 CCGCTCAGTCCCTCGGCGCAGCCCGGGGCTCGCTCCCTCGTCCCGCAACCGGAGCT 1752

1498 CCGCTCAGTCCCTCGGCGCAGCCCGGGGCTCGCTCCCTCGTCCCGCAACCGGAGCT 1557

1753 CTTTGGACCGCATCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 1812

1558 CTTTGGACCGCATCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 1617

1813 CTTTCTCTGTTGGGTTGAAATAATCAGCGGAAATCCCCCTTTCAGTTCTTTGAAAAGTT 1872

1618 CTTTCTCTGTTGGGTTGAAATAATCAGCGGAAATCCCCCTTTCAGTTCTTTGAAAAGTT 1677

1873 CCAATGACTCGAATATCTGAAATGAGAAACAAACCGACTCACAACCTCCAGTAGCTC 1932

1678 CCAATGACTCGAATATCTGAAATGAGAAACAAACCGACTCACAACCTCCAGTAGCTC 1737

1933 CAAATGCAATTTTAAATGGAATACTGAAAGAAACCTGTTAGTGGCTTTAA 1992

1738 CAAATGCAATTTTAAATGGAATACTGAAAGAAACCTGTTAGTGGCTTTAA 1797

QY 1993 GCCCAAAACGTCCTTAAGCGCTCTCGAGATGAAGACGGGGGAG-CCCCAGCCAGGT 2051
| | | | |
Db 1798 GCCCAAAACGTCCTTAAGCGCTCTCGAGATGAAGACGGGGGAGCCCCAGCCAGGT 1857
| | | | |
QY 2052 GGAGACCCCGCA-GGACGGGGGGCGCCCGGTGACCGAGGCTCCGACGCCGGCGGCC 2110
| | | | |
Db 1858 GGAGACCCCGCAGGACGCGCGGGCGCGGTGACCGAGGCTCCGACAGCCGGCGGCC 1917
| | | | |
QY 2111 TGAGGGTCGGCGCAGCAGCGTCCAAGAGGGCGCGTTTGCTCTCGGTTAAATAAG 2170
| | | | |
Db 1918 TGAGGGTCGGCGCAGCAGCGTCCAAGAGGGCGCGTTTGCTCTCGGTTAAATAAG 1977
| | | | |
QY 2171 GTTCCGTCG 2180
| | | | |
Db 1978 GTTCCGTCG 1987
| | | | |

RESULT 7
US-09-949-016-12168
; Sequence 12168, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12168
; LENGTH: 16738
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12168

Query Match 57.0%; Score 1241.6; DB 4; Length 16738;
Best Local Similarity 98.0%; Pred. No. 1.5e-303;
Matches 1331; Conservative 0; Mismatches 19; Indels 8; Gaps 7;
QY 828 GCTGTTCACGACGCGCGGAGAGAGTCTTCTCTAGAACCCAGATGAAGTGC 887
| | | | |
Db 13370 GCACCTCTTCTCCCTGCGAAGAGTCTTCTCTAGAACCCAGATGAAGTGC 13429
| | | | |
QY 888 CAGCCCTCAGAGCGGGAGCCAGCAGGCGCTGAGCAGAGCCTCCGCTGAACCTGTGACCC 947
| | | | |
Db 13430 CAGCCCTCAGAGCGGGAGCCAGCAGGCGCTGAGCAGAGCCTCCGCTGAACCTGTGACCC 13489
| | | | |
QY 948 CAGACTCCGAGCCACCCCAAGCGCGCTGGCACTGGTGTCAAAACCTCACCACCTTGGAG 1007
| | | | |
Db 13490 CAGACTCCGAGCCACCCCAAGCGCGCTGGCACTGGTGTCAAAACCTCACCACCTTGGAG 13549
| | | | |
QY 1008 CCCTGGAGCTGTGTCTCCGCCCAACCTTGTTCATATGCCGAGACCCATAGCCGCT 1067
| | | | |
Db 13550 CCCTGGAGCTGTGT-CCCCCAACCTTGTTCATATGCCGAGACCCATAGCCGCT 13608
| | | | |
QY 1068 GCAAGCGAGAGGACACAGGAGCCAGCCCTGAGTCCGACCTTGGTGGCGGGCT 1127
| | | | |
Db 13609 GCAAGCGAGAGGACACAGGAGCCAGCCCTGAGTCCGACCTTGGTGGCGGGCT 13668
| | | | |
QY 1128 GGGTCTCTCGTCCACCCGGGGGACAGACACCGGCTTGTGTCAGGCTGGGCTCTG 1187
| | | | |
Db 13669 GGGTCTCTCGTCCACCCGGGGGACAGACACCGGCTTGTGTCAGGCTGGGCTCTG 13728
| | | | |
QY 1188 TGTACCCACTCTCGGTGGTGGAGACCTTCCCTCACCCCGAGGTCTTCAAGCT 1247
| | | | |

Db 13729 TGTACCCACTCTCTGGGTGCGTGCAGACCCCTTCCCTCCACCCCGAGTCTTCCAAGCT 13788
| | | | |
QY 1248 CTGCTTCTCAGTCTTCCAAATGGAACCACTCCTCGCAGCAGCCGACTTACCAGGA 1307
| | | | |
Db 13789 CTGCTTCTCAGTCTTCCAAATGGAACCACTCCTCGCAGCAGCCGACTTACCAGGA 13848
| | | | |
QY 1308 CGCATGCCCTTCTCTCTCCCTCATCAAAACCAAGACCCGACTTCTTTCGACACC 1367
| | | | |
Db 13849 CGCATGCCCTTCTCTCTCCCTCATCAAAACCAAGACCCGACTTCTTTCGACACC 13908
| | | | |
QY 1368 CAGGCTGTGTCGGGCCCGCAGGTGTCGGGTCTCTCTCACTCCAGGGCTCCGCGCCA 1427
| | | | |
Db 13909 CAGGCTGTGTCGGGCCCGCAGGTGTCGGGTCTCTCTCACTCCAGGGCTCCGCGCCA 13968
| | | | |
QY 1428 AGTGAGGGGGCCCTTGCAGGACCTCAGACACACTGGAGTTTCAAGGC--TGCGGGGGCT 1485
| | | | |
Db 13969 AGTGAGGGGGCCCTTGCAGGACCTCAGACACACTGGAGTTTCAAGGC--TGCGGGGGCT 14028
| | | | |
QY 1486 TGGCACATACCTGTCTTGGCTATGACAGGCTTGGGGGCTTTCGCGGCGACCCCG 1545
| | | | |
Db 14029 GGCACATACCTGTCTTGGCTATGACAGGCTTGGGGGCTTTCGCGGCGACCCCG 14088
| | | | |
QY 1546 GGGCCGAGGTAGGTCTGGGGGCTTAGAGGCTGGGATGGCTTCCGCGGCCACCCGCG 1605
| | | | |
Db 14089 GGGCCGAGGTAGGTCTGGGGGCTTAGAGGCTGGGATGGCTTCCGCGGCCACCCGCG 14147
| | | | |
QY 1606 GGGCAAGCGCAGGCGCGGCTGGAGCGCGCGCGGCTGGGGCTTGGGGCTTCAAGTGG 1665
| | | | |
Db 14148 GGGC-AGCGCAGGCGCGGCTGGAGCGCGCGCGGCTGGGGCTTGGGGCTTCAAGTGG 14206
| | | | |
QY 1666 ACCTG-CTTCCGGGGCTGGTGGCGCATCTCCTCAGTCCCTCGGCCACCCGGGGCTCGCTC 1724
| | | | |
Db 14207 ACCTGCTCCACCGCAGCTTGGAGCGCTTCTTGGAGCCAGATCTGTTCATCTGTCT 1784
| | | | |
QY 14267 CCTGTGCTCCACCGCAGCTTGGAGCGCTTCTTGGAGCCAGATCTGTTCATCTGTCT 14326
| | | | |
QY 1785 TCGTCACTGCGCGGGGGCTTGTGATGCTTTCATCTGTATGGGTGGAAAAATCAACGGG 1844
| | | | |
Db 14327 TCGTCACTGCGCGGGGGCTTGTGATGCTTTCATCTGTATGGGTGGAAAAATCAACGGG 14386
| | | | |
QY 1845 AATCCCTTTCAGTCTTGTGAAAAAGTTCCATGATCGAATATCTGAAATGAAGAAAAACA 1904
| | | | |
Db 14387 AATCCCTTTCAGTCTTGTGAAAAAGTTCCATGATCGAATATCTGAAATGAAGAAAAACA 14446
| | | | |
QY 1905 AACCGACTCACAACCTCCAGTAGCTCCAAATGCAATTTTAAATGGAACAAAAAT 1964
| | | | |
Db 14447 AACCGACTCACAACCTCCAAAGTAGCTCCAAATGCAATTTTAAATGGAACAAAAAT 14506
| | | | |
QY 1965 CTGAAAGAAACGCTCTTGTAGTGGCTTAAAGCCCAAAAGCTTCCCTAAGGCGCTCTCGAGAT 2024
| | | | |
Db 14507 CTGAAAGAAACGCTCTTGTAGTGGCTTAAAGCCCAAAAGCTTCCCTAAGGCGCTCTCGAGAT 14566
| | | | |
QY 2025 GAAGACGGGGGGAG-CCCCAGCAGGTGGAGACCCCGCA-GGACGCGCGCGCGCGCT 2082
| | | | |
Db 14567 GAAGACGGGGGGAGCCCGCAGCAGGTGGAGACCCCGCAGCGAGCGCGCGCGCGCT 14626
| | | | |
QY 2083 GACCGAGGCTTCGACAGCCGCGCGCTTGGGGTGGGGCGGAGCCAGGTTCAAGAGGG 2142
| | | | |
Db 14627 GACCGAGGCTTCGACAGCCGCGCGCTTGGGGTGGGGCGGAGCCAGGTTCAAGAGGG 14686
| | | | |
QY 2143 GCGGTTTGTCTCGGGTTAAATAAGTTTCCGTCG 2180
| | | | |
Db 14687 GCGGTTTGTCTCGGGTTAAATAAGTTTCCGTCG 14724
| | | | |

RESULT 8
US-09-949-016-14678
; Sequence 14678, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

Db	14089	GGGCGCAGAGT	AGGGTC - GGGGGCTTTAGAGGCTGGAGATGGCTCCTGGCCGCCACCGCCAGG	14147
Qy	1606	GGCAAGCGCAG	CCGGGCTGGAGCGCGCGGGGCTCGGGCTCGGGGGTCAAGTGG	1665
Db	14148	GGGC - AGCG	CAGCGCGGGCTGGAGCGCGGGCGGGCTCGGGCTCGGGCTCGGGGGTCAAGTGG	14206
Qy	1666	ACGCTG - CTTCC	GGGGTGCTCGCGATCCCTCAGTCCCTCGGCACCCCGGGGTGCGCTC	1724
Db	14207	ACGCGGCCCT	CCGGGGCTGGACGCGATCCCTCAGTCCCTCGGCACCCCGGGGGTGGCTC	14266
Qy	1725	CCTGTGCCCCA	CGCACCTGCGGAGCCTCTTTCGACCCAGATCTGTTTCATGCTTTTGTCT	1784
Db	14267	CCTCGTSCCA	CCGCACTCTTCGAGCCTCTTTCGACCCAGATCTGTTTCATGCTTTTGTCT	14326
Qy	1785	TCGTCACTCG	GGGGGCCCTTTTGATGCTTTCATCTGTATGGGTGAAATAACCCCGGG	1844
Db	14327	TCGTCACTCG	GGGGGCCCTTTTGATGCTTTCATCTGTATGGGTGAAATAACCCCGGG	14386
Qy	1845	AATCCGCCCT	TCAGTTCTTTTGAAAGTTCCATGACTCGAATATCTGAATGAAGAAACA	1904
Db	14387	AATCCGCCCT	TCAGTTCTTTTGAAAGTTCCATGACTCGAATATCTGAATGAAGAAACA	14446
Qy	1905	AACCGACTCA	CAACCTCCAAGTAGCTCCAAATGCAATTTTAAATATGAAACAAAAAT	1964
Db	14447	AACCGACTCA	CAACCTCCAAGTAGCTCCAAATGCAATTTTAAATATGAAACAAAAAT	14506
Qy	1965	CTGAAAGAAA	CGTCTTTAGTGGCTTTAAGGCCAAAAAGTCCCTTAAGGCGTCTTCGAGAT	2024
Db	14507	CTGAAAGAAA	CGTCTTTAGTGGCTTTAAGGCCAAAAAGTCCCTTAAGGCGTCTTCGAGAT	14566
Qy	2025	GAAGACGGGG	GGGAG - CCCAGCCAGGTGGAGACCCCGCA - GGAAGCGGGCGGCGCCCGGT	2082
Db	14567	GAAGACGGGG	GGGAGACCCCGCAGGTGGAGACCCCGCAGGACCGCGCGGCGCCCGGT	14626
Qy	2083	GACCGAGGCT	TCGACACGCGCGCGCCCTCGAGGGTTCGGGCGGACGAGGGTCCAAAGAGG	2142
Db	14627	GACCGAGGCT	TCGACACGCGCGCGCCCTCGAGGGTTCGGGCGGACGAGGGTCCAAAGAGG	14686
Qy	2143	CGCGTTTGTG	TCGCGGTTTAAATAAGGTTCCGTCGG	2180
Db	14687	CGCGTTTGTG	TCGCGGTTTAAATAAGGTTCCGTCGG	14724

RESULT 9

US-09-949-016-26333/c
: Sequence 26333, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER,
TITLE OF INVENTION:

; TITLE OF INVENTION:	
: TITLE OF INVENTION:	

; FILE REFERENCE: CLO

3. CURRENT APPLICATION

; CURRENT FILING DATE

1. PRIOR APPLICATION NO. _____

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

PRIOR FILING DATE:

; PRIOR APPLICATION N

; PRIOR FILING DATE:
 ; NUMBER OF SEC ID NO

; NUMBER OF SEQ IS ...
; SOFTWARE: FastSEQ

; SEQ ID NO 26333

```

; LENGTH: 601

```

TYPE: DNA
ORGANISM: Human

US-09-949-016-26333

Query Match

Best Local Similarity
Matches 592: Cons

Malcolm, 1722, 1723, 1724

Qy 906 CCCAGCAG

	Matches	590;	Conservative	1;	Mismatches	7;	Indels	5;	Gaps	4;
QY	1370	GGCTGGTCCGGGCCCCACAGGTGTGGGGTCCGCTCTCTCCACTCCAGGGCTCCGGCGCCCAAG	1429							
Db	601	GGCTGGTCCGGGCCCCACAGGTGTGGGGTCCGCTCTCTCTCCACTCCAGGGCTCCGGCGCCCAAG	542							
QY	1430	TGAGGGGGGCCCTTGC CGGAGCCTCAGACACACTGGAGTTCAGAGC--TGSGGGGGGCTTG	1487							
Db	541	TGAGGGGGGCCCTTGC CGGAGCCTCAGACACACTGGAGTTCAGGGCTGTGGGGGCGCTTGG	492							
QY	1488	GCACATACCTGTCCCTTTGGCTATGAGCAGAGCTTTGGGGGCCCTTCGCGGCACGCCCGGG	1547							
Db	481	CCACATACCTGTCCCTTTGGCTATGAGCAGAGCTTTGGGGGCCCTTCGCGGCACGCCCGGG	422							
QY	1548	GGCGAGGTAGGGTCTGGGGGCTTAGAGGTGGGATGGCTCTGGGCCACGCCCAGGGG	1607							
Db	421	GGCGAGGTAGGGTCTGGGGGCTTAGAGGTGGGATGGCTCTGGGCCACGCCCAGGGG	363							
QY	1608	GCAAGCGAGGCGGGGCTGGAGCGGCGCGCGGCTCGGGGTGGGGGGTTCAGAGTGGAC	1667							
Db	362	GC-AGCGCAGGCGGGGCTGGAGCGGCGCGCGGCTCGGGGTGGGGGGTTCAGAGTGGAC	304							
QY	1668	GCTG-CCTTCGGGGGCTGGTCCGCGATCCCTCAGTCCTTCGGGCCACCCGGGGGTGCCTCCC	1726							
Db	303	GCYGCCCTTCGGGGGCTGGAGCGCGATCCCTCAGTCCTTCGGGCCACCCGGGGGTGCCTCCC	244							
QY	1727	TCGTGCCACCGGACCTTCGGAGCGCTCTTGGACCCAGATCTGTTCATGCTTTTGTCTTC	1786							
Db	243	TCGTGCCACCGGACCTTCGGAGCGCTCTTGGACCCAGATCTGTTCATGCTTTTGTCTTC	184							
QY	1787	GTCACTGGGGGGGGCCCTTTTGATGTCTTCACTGTATGGGGTGGAAAAATCACCGGAA	1848							
Db	183	GTCACTGGGGGGGGCCCTTTTGATGTCTTCACTGTATGGGGTGGAAAAATCACCGGAA	124							
QY	1847	TCCCCCTTCAGTCTTTTGAAAAAGTTCCATCACTCCGATATCTGAAATGAAGAAACAAA	1906							
Db	123	TCCCCCTTCAGTCTTTTGAAAAAGTTCCATCACTCCGATATCTGAAATGAAGAAACAAA	64							
QY	1907	CCGACTCAAAACCTCCAAAGTAGCTCCAAATGCAATTTTTTAAATTGGAAAAACAAAATCT	1966							
Db	63	CCGACTCAAAACCTCCAAAGTAGCTCCAAATGCAATTTTTTAAATTGGAAAAACAAAATCT	4							
QY	1967	GAA 1969								
Db		3 GAA 1								

```

RESULT 12
US-09-949-016-105203/C
; Sequence 105203, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105203
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
;
US-09-949-016-105203
Query Match          24.8%; Score 541.4; DB 4; Length 601;
Best Local Similarity 97.8%; Pred. No. 4.8e-127;

```

Query Match	24.8%	Score 541.4;	DB 4;	Length 601;
Best Local Similarity	97.8%	Pred. No. 4.8e-127;		

```
Query Match 7.1%; Score 155.4; DB 4; Length 1112;
Best Local Similarity 60.6%; Pred. No. 3e-29;
Matches 255; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 310 AGCCATGACAGCTGCCCTCCCTGGCATTCCCTGGCCACGTTTCCAGGCCCTTGGGACCT 369
DB 131 AGCCATGCTGGCTACTCTGTAACATCTCTGGCTGTTTCCAGATGCTCTGGGCT 190

QY 370 CTGTTTGTGGCTGCTCCTTCCAGTGTCTCAGATGAAGCTGGGACAGCCCATCTGCAC 429
DB 191 CCTTCTACTGGGGCTCCCTGAATGCCCATAAAGATGCTGGGACGAACCTTGTTCAC 250

QY 430 AGAGGGGTAGTCTCTGTGTCTTGGGGCGAGAACCCGTCATGTCTCTGCAACATCTCAA 489
DB 251 TGAGCATGAAGTATCTGTAACAGAGGCGAGCCGTGTGGTGTATGGCTGTATATCTCAA 310

QY 490 CGCCTTCTCCCATGTCAACATCAAGCTGGTCCCGGACGAGGAGGCCATCTCAA 549
DB 311 CAATCTCAGAGAGCTACCAATGAGTTGGTTACAGTGAAGACTAGCATCTTCAA 370

QY 550 TGAGTGGCTCCAGGCTACTTCTCCCGGACGGCTGGCAGCTCCAGGTTCCAGGGAGGCT 609
DB 371 TCATAGCCTCCAGGAACACTACTTAAGGATTCATGGCAGCTTCATATTCAGAGTCCA 430

QY 610 GGCACAGCTGTGATCAAGGGCGCCGGGACTCCCATGCTGGGCTGTACATGTGGCACT 669
DB 431 GGGCCAGCTGTGTATCACAGATGCTCAGGACAAACACTCAGGGAACACTACTCTGGAAGCT 490

QY 670 CGTGGACACACAGAGAAATAACAGACAGTCAAGCTGGAGGTTTCAGGTGACAGACCCA 729
DB 491 GCATGGATTCAGGAGAGTTCAAAAACCTTCAACCTGACTGTTAATGCGGCAGACAGACA 550

QY 730 G 730
DB 551 G 551

RESULT 14
US-09-949-016-26332/c
; Sequence 26332, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26332
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-26332

Query Match 6.4%; Score 138.6; DB 4; Length 601;
Best Local Similarity 99.3%; Pred. No. 4e-25;
Matches 138; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 712 TTCAGGTGCAGAACCCAGTCCGCCCTGCACACTGGGTTCTGGCTGTGCCAGCGGTGT 771
DB 328 TTCAGGTGCAGAACCCAGTCCGCCCTGCACACTGGGTTCTGGCTGTGCCAGCGGTGT 269

QY 772 CACTGCTGTCTTCATCTTCTTGGTGGCTCTGTCTATGTTCCGCTGTGTACAGTGCCTG 831
DB 268 CACTGCTGTCTTCATCTTCTTGGTGGCTCTGTCTATGTTCCGCTGTGTACAGTGCCTG 209

Search completed: March 20, 2005, 06:56:11
Job time : 386 secs
```

```
Query Match 7.1%; Score 155.4; DB 4; Length 1112;
Best Local Similarity 60.6%; Pred. No. 3e-29;
Matches 255; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 310 AGCCATGACAGCTGCCCTCCCTGGCATTCCCTGGCCACGTTTCCAGGCCCTTGGGACCT 369
DB 131 AGCCATGCTGGCTACTCTGTAACATCTCTGGCTGTTTCCAGATGCTCTGGGCT 190

QY 370 CTGTTTGTGGCTGCTCCTTCCAGTGTCTCAGATGAAGCTGGGACAGCCCATCTGCAC 429
DB 191 CCTTCTACTGGGGCTCCCTGAATGCCCATAAAGATGCTGGGACGAACCTTGTTCAC 250

QY 430 AGAGGGGTAGTCTCTGTGTCTTGGGGCGAGAACCCGTCATGTCTCTGCAACATCTCAA 489
DB 251 TGAGCATGAAGTATCTGTAACAGAGGCGAGCCGTGTGGTGTATGGCTGTATATCTCAA 310

QY 490 CGCCTTCTCCCATGTCAACATCAAGCTGGTCCCGGACGAGGAGGCCATCTCAA 549
DB 311 CAATCTCAGAGAGCTACCAATGAGTTGGTTACAGTGAAGACTAGCATCTTCAA 370

QY 550 TGAGTGGCTCCAGGCTACTTCTCCCGGACGGCTGGCAGCTCCAGGTTCCAGGGAGGCT 609
DB 371 TCATAGCCTCCAGGAACACTACTTAAGGATTCATGGCAGCTTCATATTCAGAGTCCA 430

QY 610 GGCACAGCTGTGATCAAGGGCGCCGGGACTCCCATGCTGGGCTGTACATGTGGCACT 669
DB 431 GGGCCAGCTGTGTATCACAGATGCTCAGGACAAACACTCAGGGAACACTACTCTGGAAGCT 490

QY 670 CGTGGACACACAGAGAAATAACAGACAGTCAAGCTGGAGGTTTCAGGTGACAGACCCA 729
DB 491 GCATGGATTCAGGAGAGTTCAAAAACCTTCAACCTGACTGTTAATGCGGCAGACAGACA 550

QY 730 G 730
DB 551 G 551

RESULT 14
US-09-949-016-26332/c
; Sequence 26332, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26332
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-26332

Query Match 6.4%; Score 138.6; DB 4; Length 601;
Best Local Similarity 99.3%; Pred. No. 4e-25;
Matches 138; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 712 TTCAGGTGCAGAACCCAGTCCGCCCTGCACACTGGGTTCTGGCTGTGCCAGCGGTGT 771
DB 328 TTCAGGTGCAGAACCCAGTCCGCCCTGCACACTGGGTTCTGGCTGTGCCAGCGGTGT 269

QY 772 CACTGCTGTCTTCATCTTCTTGGTGGCTCTGTCTATGTTCCGCTGTGTACAGTGCCTG 831
DB 268 CACTGCTGTCTTCATCTTCTTGGTGGCTCTGTCTATGTTCCGCTGTGTACAGTGCCTG 209
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 04:49:49 ; Search time 1256 Seconds
(without alignments)
10332.824 Million cell updates/sec

Title: US-10-080-522-2

Perfect score: 2180

Sequence: 1 ATTCTCGTCTCTTAGCGT.....TTAAATAAGTTCGTCG 2180

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 554816 seqs, 2976611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2180	100.0	2180	14	US-10-080-522-2
2	1951	89.5	2000	9	US-09-954-456-1237
3	1951	89.5	2000	9	US-09-997-165-3
4	1951	89.5	2000	18	US-10-775-169-97
5	912.8	41.9	1064	9	US-09-799-777-109
c 6	222	10.2	3026	18	US-10-676-248B-245
7	155.4	7.1	1112	9	US-09-997-165-7
8	60	2.8	60	10	US-09-908-975-11235
9	54	2.5	100301	18	US-10-450-826-83
10	54	2.5	100301	18	US-10-723-860-3574
c 11	52	2.4	606	17	US-10-424-599-95310

c 12	51.2	2.3	5452	15	US-10-017-161-1481	Sequence 1481, Ap
c 13	51.2	2.3	5452	17	US-10-292-798-1189	Sequence 1189, Ap
c 14	50.4	2.3	805	17	US-10-424-599-61903	Sequence 61903, A
c 15	50.2	2.3	671	14	US-10-184-644-346	Sequence 346, App
c 16	50.2	2.3	671	14	US-10-184-644-346	Sequence 346, App
c 17	50.2	2.3	728	18	US-10-425-115-18361	Sequence 18361, A
c 18	49.6	2.3	645	18	US-10-767-701-3807	Sequence 3807, Ap
c 19	49.6	2.3	763	18	US-10-437-963-49370	Sequence 49370, A
c 20	49.6	2.3	882	14	US-10-184-644-574	Sequence 574, App
c 21	49.6	2.3	882	14	US-10-184-644-574	Sequence 574, App
c 22	49.4	2.3	4667	18	US-10-723-860-5759	Sequence 5759, Ap
c 23	49.4	2.3	5452	15	US-10-017-161-1481	Sequence 1481, Ap
c 24	49.4	2.3	5452	17	US-10-292-798-1189	Sequence 1189, Ap
c 25	49.2	2.3	581	18	US-10-425-115-93481	Sequence 93481, A
c 26	49.2	2.3	581	18	US-10-425-115-93481	Sequence 93481, A
c 27	49.2	2.3	1506	18	US-10-437-963-38913	Sequence 38913, A
c 28	49	2.2	805	17	US-10-424-599-61903	Sequence 61903, A
c 29	48.8	2.2	959	18	US-10-437-963-48694	Sequence 48694, A
c 30	48.8	2.2	1533	18	US-10-739-930-4713	Sequence 4713, Ap
c 31	48.6	2.2	806	18	US-10-425-115-123827	Sequence 123827, A
c 32	48.6	2.2	972	18	US-10-425-115-17653	Sequence 17653, A
c 33	48.4	2.2	486	18	US-10-425-115-181872	Sequence 181872, A
c 34	48.4	2.2	737	18	US-10-425-115-65924	Sequence 65924, A
c 35	48.4	2.2	3897	17	US-10-213-796-20	Sequence 20, Appl
c 36	48.4	2.2	22210	17	US-10-213-796-19	Sequence 19, Appl
c 37	48.2	2.2	594	14	US-10-123-155-10	Sequence 10, Appl
c 38	48.2	2.2	594	15	US-10-146-731-10	Sequence 10, Appl
c 39	48.2	2.2	594	15	US-10-140-472-10	Sequence 10, Appl
c 40	48.2	2.2	594	15	US-10-141-761-10	Sequence 10, Appl
c 41	48.2	2.2	594	16	US-10-142-885-10	Sequence 10, Appl
c 42	48.2	2.2	594	16	US-10-158-790-10	Sequence 10, Appl
c 43	48.2	2.2	594	17	US-10-137-871-10	Sequence 10, Appl
c 44	48.2	2.2	594	17	US-10-140-923-10	Sequence 10, Appl
c 45	48.2	2.2	594	17	US-10-141-756-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-10-080-522-2
; Sequence 2, Application US/10080522
; Publication No. US20030096326A1
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/080,522
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/539,774
; FILING DATE: 31-MAR-2000
; APPLICATION NUMBER: US 09/210,474
; FILING DATE: 14-DEC-1998
; APPLICATION NUMBER: US 08/755,559
; FILING DATE: 22-NOV-1996
; ATTORNEY/AGENT INFORMATION:

[illegible]

Db 1861 TTGTGAAAGTTCCATGACTCGAATATCTGAATGAAGAAAACAAACCGACTCACAAACC 1920

Qy 1921 TCCAGTAGCTCCAAATGCAATTTTTTAAATGGAACAAAATCTGAAGAAACGTCCTT 1980

Db 1921 TCCAGTAGCTCCAAATGCAATTTTTTAAATGGAACAAAATCTGAAGAAACGTCCTT 1980

Qy 1981 TAGTGGCTTTAAGCCCAAAACGTCCTTAAGCGGTCTCTCGAGATGAAGACGGGGGGAGC 2040

Db 1981 TAGTGGCTTTAAGCCCAAAACGTCCTTAAGCGGTCTCTCGAGATGAAGACGGGGGGAGC 2040

Qy 2041 CCAGCGAGGTGAGACCCCGCAGAGACGGCGGGCGCGCGCGGTGACCGAGGCCCTCGACAG 2100

Db 2041 CCAGCGAGGTGAGACCCCGCAGAGACGGCGGGCGCGCGGTGACCGAGGCCCTCGACAG 2100

Qy 2101 CCGGCGCGCTTGAAGGTCTGGGCGCGAGCCAGAGGTTCACAGAGGGCGCGTTTGTCTCGGG 2160

Db 2101 CCGGCGCGCTTGAAGGTCTGGGCGCGAGCCAGAGGTTCACAGAGGGCGCGTTTGTCTCGGG 2160

Qy 2161 TTAAATAAAGTTCCGTCGC 2180

Db 2161 TTAAATAAAGTTCCGTCGC 2180

RESULT 2

US-09-954-456-1237

; Sequence 1237, Application US/09954456

; Patent No. US20020115057A1

; GENERAL INFORMATION:

; APPLICANT: Young, Paul

; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand

; TITLE OF INVENTION: Sets

; FILE REFERENCE: 689290-76

; CURRENT APPLICATION NUMBER: US/09/954,456

; PRIORITY FILING DATE: 2001-09-18

; PRIORITY APPLICATION NUMBER: US/60/233,617

; PRIORITY FILING DATE: 2000-09-18

; PRIORITY APPLICATION NUMBER: US/60/234,052

; PRIORITY FILING DATE: 2000-09-20

; PRIORITY APPLICATION NUMBER: US/60/234,923

; PRIORITY FILING DATE: 2000-09-25

; PRIORITY APPLICATION NUMBER: US/60/235,134

; PRIORITY FILING DATE: 2000-09-25

; PRIORITY APPLICATION NUMBER: US/60/235,637

; PRIORITY FILING DATE: 2000-09-26

; PRIORITY APPLICATION NUMBER: US/60/235,638

; PRIORITY FILING DATE: 2000-09-26

; PRIORITY APPLICATION NUMBER: US/60/235,711

; PRIORITY FILING DATE: 2000-09-27

; PRIORITY APPLICATION NUMBER: US/60/235,720

; PRIORITY FILING DATE: 2000-09-27

; PRIORITY APPLICATION NUMBER: US/60/235,840

; PRIORITY FILING DATE: 2000-09-27

; PRIORITY APPLICATION NUMBER: US/60/235,863

; PRIORITY FILING DATE: 2000-09-27

; NUMBER OF SEQ ID NOS: 2276

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1237

; LENGTH: 2000

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-954-456-1237

Query Match 89.5%; Score 1951; DB 9; Length 2000;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

Qy 196 ATTTTCTGGGCTCCGGGCGGAGAAAGCTGCATCCAGAGGAGCGGTCCAGAGCG 255

Db 1 ATTTTCTGGGCTCCGGGCGGAGAAAGCTGCATCCAGAGGAGCGGTCCAGAGCG 60

Qy 256 GACCCGGGAGTGTTCAGAGCCAGTGAAGGACCGAGGCGCCCAAGTCCGACCCACCCAT 315

Db 61 GACCCGGGAGTGTTCAGAGCCAGTGAAGGACCGAGGCGCCCAAGTCCGACCCACCCAT 120

Qy 316 GCAGACCTGCCCCCTTGGCATTTCCCTGGCCACGTTTTCCAGGCCCTTGGAGCCCTCTGTT 375

Db 121 GCAGACCTGCCCCCTTGGCATTTCCCTGGCCACGTTTTCCAGGCCCTTGGAGCCCTCTGTT 180

Qy 376 TTTGGCTGCTCTTGGCTGCTCAGATGAAGGCTGGGACAGCCCATCTGCACAGAGG 435

Db 181 TTTGGCTGCTCTTGGCTGCTCAGATGAAGGCTGGGACAGCCCATCTGCACAGAGG 240

Qy 436 GGTAGTCTCTGTCTTGGGGCGAACAACCGTCAATGCTCTCAACATCTCCAAAGCCCTT 495

Db 241 GGTAGTCTCTGTCTTGGGGCGAACAACCGTCAATGCTCTCAACATCTCCAAAGCCCTT 300

Qy 496 TTCCCATGTCAACATCAAGCTGCTGCCCAAGGCGAGAGCGCCATCTTCAATGAGGT 555

Db 301 TTCCCATGTCAACATCAAGCTGCTGCCCAAGGCGAGAGCGCCATCTTCAATGAGGT 360

Qy 556 GGTCCAGGCTACTTCTCCCGGAGCGCTGCCAGCTCCAGGTTCAGGAGGCGGTGGCACA 615

Db 361 GGTCCAGGCTACTTCTCCCGGAGCGCTGCCAGCTCCAGGTTCAGGAGGCGGTGGCACA 420

Qy 616 GCTGGTGATCAAGGCGCGGGACTCCCATGCTGGGCTGTACATGTGGACCTCTGTTGG 675

Db 421 GCTGGTGATCAAGGCGCGGGACTCCCATGCTGGGCTGTACATGTGGACCTCTGTTGG 480

Qy 676 ACACAGAGAAATAACAGACAAGTCAAGTCAAGTGGAGGTTTTAGGTGCAGAACCCAGTCCGC 735

Db 481 ACACAGAGAAATAACAGACAAGTCAAGTCAAGTGGAGTTTTAGGTGCAGAACCCAGTCCGC 540

Qy 736 CCTGACACTGGGTTCTGGCTGTGCCAGCGGTGCTCACTGTCTTTCATCTCTTGGT 795

Db 541 CCTGACACTGGGTTCTGGCTGTGCCAGCGGTGCTCACTGTCTTTCATCTCTTGGT 600

Qy 796 CGCTCTGGTCAATGTCGCTGTCAGGTGCGCGCTGTTCCAGCAACCGCGGAGAGAA 855

Db 601 CGCTCTGGTCAATGTCGCTGTCAGGTGCGCGCTGTTCCAGCAACCGCGGAGAGAA 660

Qy 856 GTTCTTCTCTCTAGAACCCAGATGAAGTGCAGCCCTCAGAGCGGAGCCAGCAGGG 915

Db 661 GTTCTTCTCTCTAGAACCCAGATGAAGTGCAGCCCTCAGAGCGGAGCCAGCAGGG 720

Qy 916 CCTGAGCAGAGGCTCCGCTGAACTGTGGACCCAGACTCCGAGCCCAACCCAGGCCGT 975

Db 721 CCTGAGCAGAGGCTCCGCTGAACTGTGGACCCAGACTCCGAGCCCAACCCAGGCCGT 780

Qy 976 GGCAGTGTGTTTCAAAACCTCACCCTTGGAGCCCTGAGCTGCTGCTCCGCCCAACCT 1035

Db 781 GGCAGTGTGTTTCAAAACCTCACCCTTGGAGCCCTGAGCTGCTGCT -CCCCCACAACCT 839

Qy 1036 TGTTCATATGCGCAGACCCCATAGCCGCTTGAAGGCGAGAGGACACAGAGAGCCA 1095

Db 840 TGTTCATATGCGCAGACCCCATAGCCGCTTGAAGGCGAGAGGACACACAGAGAGCCA 899

Qy 1096 GGCCTGAGTGGGACCTTGGGTGGGGGGCTGGGTCTCTGCTCCAGCCCGAGGGGACA 1155

Db 900 GGCCTGAGTGGGACCTTGGGTGGGGGGCTGGGTCTCTGCTCCAGCCCGAGGGGACA 959

Qy 1156 GACACCGGCTTGTGGCAGGCTGGGCTCTGTGTACCCACTCTCTGGGTGCTGCAGAC 1215

Db 960 GACACCGGCTTGTGGCAGGCTGGGCTCTGTGTACCCACTCTCTGGGTGCTGCAGAC 1019

Qy 1216 CTTTCCCTTCCACCCCGCAGGCTTTCAGAGCTCTCTTCTCAGTTTTCCAAATGGAACC 1275

Db 1020 CTTTCCCTTCCACCCCGCAGGCTTTCAGAGCTCTCTTCTCAGTTTTCCAAATGGAACC 1079

Qy 1276 ACCTCACTCCGAGCAGCCGACTTACAGAGCGATGCCCCCTCTCTGCTCTCATCA 1335

Db 1080 ACCTCACTCCGAGCAGCCGACTTACAGAGCGATGCCCCCTCTCTGCTCTCATCA 1139

Qy 1336 ACCCAGACCGCGGACTCCCTTTCTGCGACCCAGCGTGTGCTGGGCTGGGCT 1395

Db 1140 ACCCAGACCGCGGACTCCCTTTCTGCGACCCAGCGTGTGCTGGGCTGGGCT 1199

1396 CCGCTCTCTCCACTCCCGGGCTCCGCGCCCAAGTGAAGGGGGCCCTGCGGAGCCTCAG 1455
1200 CCGCTCTCTCCACTCCCGGGCTCCGCGCCCAAGTGAAGGGGGCCCTGCGGAGCCTCAG 1259
1456 ACACATGAGTTAGGCTGGGGGGCTTGGGCATACCTGCTCCCTTGGCTATGACA 1515
1260 ACACATGAGTTAGGCTGGGGGGCTTGGGCATACCTGCTCCCTTGGCTATGACA 1319
1516 GCTTTGGGGGCCCTTCCGCGGACGCGCGGGGCGGAGGTAGGTCTGGGGGCTTAGAG 1575
1320 GCTTTGGGGGCCCTTCCGCGGACGCGCGGGGCGGAGGTAGGTCTGGGGGCTTAGAG 1379
1576 GCTTGGATGGCTCTTGGGCCCAACGCGAGGGGGCAAGCGAGGCTGGAGGGCGGC 1635
1380 GCTTGGATGGCTCTTGGGCCCAACGCGAGGGGGCAAGCGAGGCTGGAGGGCGGC 1439
1636 GCGGGGGCTCGGGCTGGGGGCTCAGGTGAGCGCTCCGCGGCTGGTCCGCGATCCC 1695
1440 GCGGGGGCTCGGGCTGGGGGCTCAGGTGAGCGCTCCGCGGCTGGTCCGCGATCCC 1499
1696 TCAGTCCCTCGGCGACCCGCGGGTCCCTCTGTCGCCACCGCACCTGCGAGCCTCTT 1755
1500 TCAGTCCCTCGGCGACCCGCGGGTCCCTCTGTCGCCACCGCACCTGCGAGCCTCTT 1559
1756 TGAACCAAGATCTGTTTCTTCTTCTGCTCACTGGCGGGGCGCTTTGATGCTT 1815
1560 TGAACCAAGATCTGTTTCTTCTTCTGCTCACTGGCGGGGCGCTTTGATGCTT 1619
1816 CATCTGATGGGTGGAATAATCACCGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCA 1875
1620 CATCTGATGGGTGGAATAATCACCGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCA 1679
1876 TGACTCGAATATCTGAATGAAGAAACAAACCGAATCAAACTCCAAAGTAGTCCAA 1935
1680 TGACTCGAATATCTGAATGAAGAAACAAACCGAATCAAACTCCAAAGTAGTCCAA 1739
1936 ATGCAATTTTAAATGGAACAAATACTGAAGAAAGCTCTTTAGTGGCTTTAAGCC 1995
1740 ATGCAATTTTAAATGGAACAAATACTGAAGAAAGCTCTTTAGTGGCTTTAAGCC 1799
1996 CCAAAAGCTCCCTAAGCGCTCCTCGAGATGAAGACGCGGGGGAG - CCGCAGCGAGTGA 2054
1800 CCAAAAGCTCCCTAAGCGCTCCTCGAGATGAAGACGCGGGGGAGCCCGCAGGTGA 1859
2055 GACCCCGAGGACGCGCGGGCGCGCTGACCGAGCGCTCGACGCGCGCGCTGAG 2114
1860 GACCCCGAGGACGCGCGGGCGCGCTGACCGAGCGCTCGACGCGCGCGCTGAG 1919
2115 GGTGCGGC - GAGCAGCGTCCAGAGGGGCGCGCTTTGCTCTCGGTTAAAAATAAGTT 2173
1920 GGTGCGGCAGGACGAGGTCCAGAGGGGCGCGTTTGTCTCGGTTAAAAATAAGTT 1979
2174 CCGTCCG 2180
1980 CCGTCCG 1986

RESULT 3
US-09-997-165-3
; Sequence 3, Application US/09997165
; Patent No. US20020141999A1
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fanslow, William C.
; TITLE OF INVENTION: LGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(865)
US-09-997-165-3

Query Match 89.5%; Score 1951; DB 9; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
Qy 196 ATTTCTCTGGGGCTCCGGGGCGCGGAGAGCTCATCCAGAGAGAGCGGTCCAGAGCG 255
Db 1 ATTTCTCTGGGGCTCCGGGGCGCGGAGAGCTCATCCAGAGAGAGCGGTCCAGAGCG 60
Qy 256 GACCCGGGAGTGTTCAGAGAGCCAGTGAACAAGACCAAGGGGCCCAAGTCCACAGCCAT 315
Db 61 GACCCGGGAGTGTTCAGAGAGCCAGTGAACAAGACCAAGGGGCCCAAGTCCACAGCCAT 120
Qy 316 GCAGACCTGCCCTGGCATTCCTGGCCACCTTTCCAGGCCCTTGGGACCCCTCTCTGT 375
Db 121 GCAGACCTGCCCTGGCATTCCTGGCCACCTTTCCAGGCCCTTGGGACCCCTCTCTGT 180
Qy 376 TTTGGCTGCCCTTGAAGTGTCTCAGAAATGAAGGTGGGACAGCCCAATCTGACAGAGG 435
Db 181 TTTGGCTGCCCTTGAAGTGTCTCAGAAATGAAGGTGGGACAGCCCAATCTGACAGAGG 240
Qy 436 GGTAGTCTGTGTCTTGGGGCGAGAACACCGTCATGCTCTGCAACATCTCCAACGCTT 495
Db 241 GGTAGTCTGTGTCTTGGGGCGAGAACACCGTCATGCTCTGCAACATCTCCAACGCTT 300
Qy 496 CTCCATGTCAACATCAAGTCTGCGTCCAGCGGACAGAGAGCGCCATCTTCAATGAGGT 555
Db 301 CTCCATGTCAACATCAAGTCTGCGTCCAGCGGACAGAGAGCGCCATCTTCAATGAGGT 360
Qy 556 GGTCTCAGGCTACTTCTCCGGGACGCTGCGAGTCAAGTTCAGGGTTCAGGGAGCGTGGACA 615
Db 361 GGTCTCAGGCTACTTCTCCGGGACGCTGCGAGTCAAGTTCAGGGTTCAGGGAGCGTGGACA 420
Qy 616 GGTGTGTATCAAGGGCGCCGGGACTCCATGCTGGGTGTACATGTGGCACTCTGTGGG 675
Db 421 GGTGTGTATCAAGGGCGCCGGGACTCCATGCTGGGTGTACATGTGGCACTCTGTGGG 480
Qy 676 ACACAGAGAAATAACAGCAAGTCAAGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTC 735
Db 481 ACACAGAGAAATAACAGCAAGTCAAGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTC 540
Qy 736 CCCTGACATCGGGTCTGCGCTGTCAGCGGTGTCTGCTGCTCTCTCTCTCTCTCTCT 795
Db 541 CCCTGACATCGGGTCTGCGCTGTCAGCGGTGTCTGCTGCTCTCTCTCTCTCTCTCT 600
Qy 796 CGCTCTGGTTCATGTTCCGCTGTCAGAGTTCGCGCTGTTCCAGCAACCGCGGAGAGAA 855
Db 601 CGCTCTGGTTCATGTTCCGCTGTCAGAGTTCGCGCTGTTCCAGCAACCGCGGAGAGAA 660
Qy 856 GTTCTTCTCTTAGAACCCCATAGAGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTC 915
Db 661 GTTCTTCTCTTAGAACCCCATAGAGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTC 720
Qy 916 CCTGAGCAGAGCTCGCTGTAAGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTCAG 975
Db 721 CCTGAGCAGAGCTCGCTGTAAGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTCAG 780
Qy 976 GGCACCTGGTTCACAAACCCCTCACTTGGAGCCCTTGGAGCTGTGTCTCTCTCTCTCT 1035
Db 781 GGCACCTGGTTCACAAACCCCTCACTTGGAGCCCTTGGAGCTGTGTCTCTCTCTCTCT 839
Qy 1036 TGTTCATATGCGCAGACCCCATAGAGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTC 1095
Db 840 TGTTCATATGCGCAGACCCCATAGAGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTC 899

1096 GGCCTGAGTGCAGACCTTGGGTGGCGGGGCTCTCTGTCACCACTCTCTGGGTGGTGCAGAC 1155
Db GGCCTGAGTGCAGACCTTGGGTGGCGGGGCTCTCTGTCACCACTCTCTGGGTGGTGCAGAC 959
1156 GACACCGGCTTGGTGGCAGGCTGGGCTCTGTGTACCACTCTCTGGGTGGTGCAGAC 1215
Db GACACCGGCTTGGTGGCAGGCTGGGCTCTGTGTACCACTCTCTGGGTGGTGCAGAC 1019
1216 CTTTCCCTCCACCCGCCAGGCTTCCAAAGCTCTGTCTCTCAGTTTCCAAATGAACC 1275
Db CTTTCCCTCCACCCGCCAGGCTTCCAAAGCTCTGTCTCTCAGTTTCCAAATGAACC 1079
1276 ACTCAGCTCCGAGACCCGACTTACAGAGAGGATGCCCTCTGCTCTGCTCATCAA 1335
Db ACTCAGCTCCGAGACCCGACTTACAGAGAGGATGCCCTCTGCTCTGCTCATCAA 1139
1336 ACCCAGACCCGAGCTCTTCTGACACCCAGGCTGGTCCGGGCCAGGCTGGGT 1395
Db ACCCAGACCCGAGCTCTTCTGACACCCAGGCTGGTCCGGGCCAGGCTGGGT 1199
1396 CGGCTCTCTCACTCCAGGCTCCGCGCCCAAGTGAAGGGGCCCTGCGGAGCTCAG 1455
Db CGGCTCTCTCACTCCAGGCTCCGCGCCCAAGTGAAGGGGCCCTGCGGAGCTCAG 1259
1456 ACACATGAGTTCAGGGCTGGGGGCTTGGCACATACCTGTCTCTGGCTATGAGCA 1515
Db ACACATGAGTTCAGGGCTGGGGGCTTGGCACATACCTGTCTCTGGCTATGAGCA 1319
1516 GGCCTTGGGGCTTCCGCGCAGCCCGGGGCGAGGTAGGTCTGGGGCTTAGAG 1575
Db GGCCTTGGGGCTTCCGCGCAGCCCGGGGCGAGGTAGGTCTGGGGCTTAGAG 1379
1576 GCTGGATGGCTCTTGGGCCACCGCAGGGGCAAGCGAGCGCGGCTGGAGGCGC 1635
Db GCTGGATGGCTCTTGGGCCACCGCAGGGGCAAGCGAGCGCGGCTGGAGGCGC 1439
1636 GCGCGGCTCGGGCTGGGGGTCAGGTGACCGCTCCGCGGCTGGTCCGCGATCCC 1695
Db GCGCGGCTCGGGCTGGGGGTCAGGTGACCGCTCCGCGGCTGGTCCGCGATCCC 1499
1696 TCAGTCCCTCGGCCACCGGGGGTGGCTCTCTGTCACCTGCGCCACCGACCTGCGAGCTCTT 1755
Db TCAGTCCCTCGGCCACCGGGGGTGGCTCTCTGTCACCTGCGCCACCGACCTGCGAGCTCTT 1559
1756 TGGACCCAGATCTGTTTCTGTTTGTCTTCTGTCACCTGCGGGGCGGCTTGTATGCTT 1815
Db TGGACCCAGATCTGTTTCTGTTTGTCTTCTGTCACCTGCGGGGCGGCTTGTATGCTT 1619
1816 CATCTGATGGGTGAAAAATCACCGGGAATCCCCCTTCACTTTTGAAGTTCCA 1875
Db CATCTGATGGGTGAAAAATCACCGGGAATCCCCCTTCACTTTTGAAGTTCCA 1679
1876 TGACTGGAATCTGAATGAAGAAAACAAACCGACTCAAACTCCAGTGTCCAA 1935
Db TGACTGGAATCTGAATGAAGAAAACAAACCGACTCAAACTCCAGTGTCCAA 1739
1936 ATGCAATTTTAAATGAAAAACAAATCTGAAAGAACGCTTTTGTGGCTTTAAGCC 1995
Db ATGCAATTTTAAATGAAAAACAAATCTGAAAGAACGCTTTTGTGGCTTTAAGCC 1799
1996 CAAAAACGTCCTAAGGGCTCTCGAGATGAAGAGCGGGGGGAG-CCCCAGCAGGTGGA 2054
Db CAAAAACGTCCTAAGGGCTCTCGAGATGAAGAGCGGGGGGAG-CCCCAGCAGGTGGA 1859
2055 GACCCGAGAGACCGGGCGGCGGCTGACCGAGGCTTCGACAGCGGCGGCGGCTGAG 2114
Db GACCCGAGAGACCGGGCGGCGGCTGACCGAGGCTTCGACAGCGGCGGCGGCTGAG 1919
2115 GGTCTGGGGC-GAGCCAGAGGTCAAGAGGGGCGGCTTGTCTCGGGTTAAATAGGTT 2173
Db GGTCTGGGGC-GAGCCAGAGGTCAAGAGGGGCGGCTTGTCTCGGGTTAAATAGGTT 1979

QY 2174 CCGTCCG 2180
Db 1980 CCGTCCG 1986
RESULT 4
US-10-775-169-97
; Sequence 97, Application US/10775169
; Publication No. US20040175743A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dornier, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 97
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-775-169-97
Query Match 89.5%; Score 1951; DB 18; Length 2000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
QY 196 ATTTTCTCGGGCTCCGGGGCGGAGAGCTGCATCCAGAGGAGCGCTCCAGAGCG 255
Db 1 ATTTTCTCGGGCTCCGGGGCGGAGAGCTGCATCCAGAGGAGCGCTCCAGAGCG 60
QY 256 GACCCGGAGTGTTCAGAGCAGTGAAGACAGAGGAGGAGGAGGAGGAGGAGGAGG 315
Db 61 GACCCGGAGTGTTCAGAGCAGTGAAGACAGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 316 GACAGACCTGCCCCCTGGGCAATCCCTGGGCGGAGGAGGAGGAGGAGGAGGAGG 375
Db 121 GACAGACCTGCCCCCTGGGCAATCCCTGGGCGGAGGAGGAGGAGGAGGAGGAGG 180
QY 376 TTTGGCTGCTCTTGGTGTCTCAGATGAAGGCTGGGAGAGGAGGAGGAGGAGGAGG 435
Db 181 TTTGGCTGCTCTTGGTGTCTCAGATGAAGGCTGGGAGAGGAGGAGGAGGAGGAGG 240
QY 436 GGTAGTCTGTGTCTTGGGGGAGAACACCGTCTATGTCTGCAACATCTCCAGGCTT 495
Db 241 GGTAGTCTGTGTCTTGGGGGAGAACACCGTCTATGTCTGCAACATCTCCAGGCTT 300
QY 496 CTCCCATGTCAACATCAAGCTGCGTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 555
Db 301 CTCCCATGTCAACATCAAGCTGCGTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 556 GGTCTCAGGCTACTTCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 615
Db 361 GGTCTCAGGCTACTTCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 616 GCTGTGTATCAAGGCGCCCGGAGCTCCCATGTCTGGGTGTATGTGTGGGAGCTCTGG 675
Db 421 GCTGTGTATCAAGGCGCCCGGAGCTCCCATGTCTGGGTGTATGTGTGGGAGCTCTGG 480
QY 676 ACACAGAGAAATACAGACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 735
Db 481 ACACAGAGAAATACAGACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 540
QY 736 CCCTGACACTGGGTCTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 795
Db 541 CCCTGACACTGGGTCTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 600
QY 796 CGCTCTGT 855

Db	601	CGCTCTGGTTCATGTTTCGCTGTCAGAGTGGCGCTGTTTCCAGCAAACGCCGGAGAGAA	660
QY	856	GTTCCTCTCTAGAACCCAGATCAAGTTCGACGCTCTCAGAGCGGAGCCACGACGGG	915
Db	661	GTTCCTCTCTAGAACCCAGATCAAGTTCGACGCTCTCAGAGCGGAGCCACGACGGG	720
QY	916	CCTGAGCAGAGCTCCGCTGAACTCTGACACCCAGACTCCGAGCCACCCCAAGCCGCT	975
Db	721	CCTGAGCAGAGCTCCGCTGAACTCTGACACCCAGACTCCGAGCCACCCCAAGCCGCT	780
QY	976	GGCACTGGTGTTCAAACCCCTACCACTTGGAGCCCTGGAGTGTCTGTCCTCCCCCAACCT	1035
Db	781	GGCACTGGTGTTCAAACCCCTACCACTTGGAGCCCTGGAGTGTCTGTCCTCCCCCAACCT	839
QY	1036	TGTTTCCATATCCCGCAGACCCATAGCCGCTGCAAGGCAGAGAGGACACAGGAGCCCA	1095
Db	840	TGTTTCCATATCCCGCAGACCCATAGCCGCTGCAAGGCAGAGAGGACACAGGAGCCCA	899
QY	1096	GGCCTGAGTGGCGACTTGGGTGGCGGGCTGGGTCTCTCTGCCACCCCGAGGGCACA	1155
Db	900	GGCCTGAGTGGCGACTTGGGTGGCGGGCTGGGTCTCTCTGCCACCCCGAGGGCACA	959
QY	1156	GACACCGGCTTCTTGGCAGGCTGGGCTCTGTGTACCCACTCTCTGGGTGGTGCAGAC	1215
Db	960	GACACCGGCTTCTTGGCAGGCTGGGCTCTGTGTACCCACTCTCTGGGTGGTGCAGAC	1019
QY	1216	CTTTCCTCTCAGCCCGCAGGCTTCCAGCTCTGTTCTCTCAGTTTCCAAATGGAACC	1275
Db	1020	CTTTCCTCTCAGCCCGCAGGCTTCCAGCTCTGTTCTCTCAGTTTCCAAATGGAACC	1079
QY	1276	ACTCACTCCGAGCACCAGCTTACAGAGCAGATGCCCTCTCTGCCCTCATCAA	1335
Db	1080	ACTCACTCCGAGCACCAGCTTACAGAGCAGATGCCCTCTCTGCCCTCATCAA	1139
QY	1336	ACCACAGACCCGAGCTTCTTCTGCCACCCAGGCTGGTCCGGCCCGAGGTGGGT	1395
Db	1140	ACCACAGACCCGAGCTTCTTCTGCCACCCAGGCTGGTCCGGCCCGAGGTGGGT	1199
QY	1396	CGGCTCTCTCACTCCAGGGCTCCGCGCCAGTGAAGGGGCGCCCTCCGAGGCTCAG	1455
Db	1200	CGGCTCTCTCACTCCAGGGCTCCGCGCCAGTGAAGGGGCGCCCTCCGAGGCTCAG	1259
QY	1456	ACACACTGGAGTTCCAGGGCTGGGGGGCTTGGCACAATCTGTCTCCCTTGGCTATGACA	1515
Db	1260	ACACACTGGAGTTCCAGGGCTGGGGGGCTTGGCACAATCTGTCTCCCTTGGCTATGACA	1319
QY	1516	GGCTTTGGGGGCTTCCGCGGAGCCCGGGGCGAGGTAGGTCTGGGGCTTAGAG	1575
Db	1320	GGCTTTGGGGGCTTCCGCGGAGCCCGGGGCGAGGTAGGTCTGGGGCTTAGAG	1379
QY	1576	GCTGGATGGCTCTGCGCCCAAGCGAGGGGCAAGCGAGGCGGGCTGGAGGCGGC	1635
Db	1380	GCTGGATGGCTCTGCGCCCAAGCGAGGGGCAAGCGAGGCGGGCTGGAGGCGGC	1439
QY	1636	GGCGGGGCTCGGGCTGGGGGCTCAGGTGGACCTGCTCCCGGGGCTGGTTCGCGCATCCC	1695
Db	1440	GGCGGGGCTCGGGCTGGGGGCTCAGGTGGACCTGCTCCCGGGGCTGGTTCGCGCATCCC	1499
QY	1696	TCAGTCCCTCGGCAACCCGGGGTTCGCTCCCTCGTCCCAACCGACCTGCGAGCTCTT	1755
Db	1500	TCAGTCCCTCGGCAACCCGGGGTTCGCTCCCTCGTCCCAACCGACCTGCGAGCTCTT	1559
QY	1756	TGGACCCAGATCTGTCATGCTTTGCTGTCGTCAGTGGGGGCGCTTTCATGCTT	1815
Db	1560	TGGACCCAGATCTGTCATGCTTTGCTGTCGTCAGTGGGGGCGCTTTCATGCTT	1619
QY	1816	CATCTGTATGGGTGGAAAAATCACCGGGAATCCCTTCCAGTCTTTGAAAAAGTTCCA	1875
Db	1620	CATCTGTATGGGTGGAAAAATCACCGGGAATCCCTTCCAGTCTTTGAAAAAGTTCCA	1679
QY	1876	TGACTCGAATATCTGAATGAAGAAACAAACCGACTCACAACCTCCAGTAGTCCCA	1935
Db	1680	TGACTCGAATATCTGAATGAAGAAACAAACCGACTCACAACCTCCAGTAGTCCCA	1739

RESULT 5
US-09-799-777-109
; Sequence 109, Application US/09799777
; Patent No. US20020091244A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Hillman, Jennifer L.
; Corley, Neil C.
; Guegler, Karl J.
; Baugh, Mariah
; Sather, Susan
; Shah, Purvi
; TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS
; NUMBER OF SEQUENCES: 154
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/799,777
; FILING DATE: 06-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/002,485
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, LUCY J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0459 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 845-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1064 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 109 :
US-09-799-777-109
Query Match 41.9%; Score 912.8; DB 9; Length 1064;
Best Local Similarity 96.1%; Pred. No. 1.4e-254;

Matches 957; Conservative 0; Mismatches 37; Indels 2; Gaps 2;
QY 199 TTCTCTGGGCTCCGGGCGCGAGAGCTGCATCCAGAGGAGCGGTCCAGAGCGGAC 258
Db 1 TTCTCTGGGCTCCGGGCGCGAGAGCTGCATCCAGAGGAGCGGTCCAGAGCGGAC 60
QY 259 CCGGAGTGTTCAGAGCCAGTGACAGGACACAGGGGCCCAAGTCCACAGCCATGCA 318
Db 61 CCGGAGTGTTCAGAGCCAGTGACAGGACACAGGGGCCCAAGTCCACAGCCATGCA 120
QY 319 GACCTGCCCCCTTGGCAATCCCTGGCCACGCTTTCCAGGGCCCTTGGAGCCCTCTGTGTTT 378
Db 121 GACCTGCCCCCTTGGCAATCCCTGGCCACGCTTTCCAGGGCCCTTGGAGCCCTCTGTGTTT 180
QY 379 GGCTGCTCTTGGAGTGTTCAGATGAAGGCTGGGACAGCCCATCTGACAGAGGGGT 438
Db 181 GGCTGCTCTTGGAGTGTTCAGATGAAGGCTGGGACAGCCCATCTGACAGAGGGGT 240
QY 439 AGTCTCTGTGCTTGGGGGAGAACACCGTCAATGCTCTGCAACATCTCCAGCCCTTCTC 498
Db 241 AGTCTCTGTGCTTGGGGGAGAACACCGTCAATGCTCTGCAACATCTCCAGCCCTTCTC 300
QY 499 CCATGTCAACATCAAGCTCGTGGCCACAGGAGAGCGGCATCTTCAATGAGTGGC 558
Db 301 CCATGTCAACATCAAGCTCGTGGCCACAGGAGAGCGGCATCTTCAATGAGTGGC 360
QY 559 TCCAGGCTACTTCTCCGGGACCGGTGGCAGTCCAGGTTTCAGGAGGCGTGGCAGGT 618
Db 361 TCCAGGCTACTTCTCCGGGACCGGTGGCAGTCCAGGTTTCAGGAGGCGTGGCAGGT 420
QY 619 GGTGATCAAGGGCCCGGACTCCATGCTGGCTGTATGTTGGACCTCTGTTGGGACA 678
Db 421 GGTGATCAAGGGCCCGGACTCCATGCTGGCTGTATGTTGGACCTCTGTTGGGACA 480
QY 679 CCAGAGAAATACAGACAAGTCAAGTGGAGTTTCAGTGAGAGAACCCAGTCCGCCCC 738
Db 481 CCAGAGAAATACAGACAAGTCAAGTGGAGTTTCAGTGAGAGAACCCAGTCCGCCCC 540
QY 739 TGACACTGGGTTCTGGCTGTGCTGACAGGCTGTGCTGCTGCTGCTGCTGCTGCTG 798
Db 541 CGACACTGGGTTCTGGCTGTGCTGACAGGCTGTGCTGCTGCTGCTGCTGCTGCTG 600
QY 799 TCTGTCATGTTGGCTGTGACAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 858
Db 601 TCTGTCATGTTGGCTGTGACAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 859 CTTCCTCTAGAACCCAGATGAAGTCCAGCCCTCAGAGCGGAGCCAGCAGGGCT 918
Db 661 CTTCCTCTAGAACCCAGATGAAGTCCAGCCCTCAGAGCGGAGCCAGCAGGGCT 720
QY 919 GAGCAGAGCTCCGCTGAACCTGTGAGCCCGCAGACTCCGAGCCACCCCAAGGCCGCTGGC 978
Db 721 GAGCAGAGCTCCGCTGAACCTGTGAGCCCGCAGACTCCGAGCCACCCCAAGGCCGCTGGC 780
QY 979 ACTGTTGTTCAACCTCACCATTGGAGCCCTTGAGCTGCTGCTGCTGCTGCTGCTGCT 1038
Db 781 ACTGTTGTTCAACCTCACCATTGGAGCCCTTGAGCTGCTGCTGCTGCTGCTGCTGCT 839
QY 1039 TTCCATATGCGCAGACCATAGCCCTGCAAGGAGAGGAGGACACAGGAGGAGCCAGCC 1098
Db 840 TTCCATATGCGCAGACCATAGCCCTGCAAGGAGAGGAGGACACAGGAGGAGGAGCCAGCC 899
QY 1099 CTGAGTCCGACCTTGGGTGGC-GGGGCTGGGTCTCTGCTCCCAAGGAGGAGGACAGA 1157
Db 900 CTGAGTCCGACCTTGGGTGGGCGGGCTGGGTCTCTGCTCCCAAGGAGGAGGACAGA 959
QY 1158 CACCGGCTTGTGGAGGCTGGGCTCTGTGTAC 1193
Db 960 AGACACCGGGCTTGTGGCAAGGCTGGGGCCCTC 995

RESULT 6

US-10-676-248B-245/c

; Sequence 245, Application US/10676248B
; Publication No. US20040161773A1
; GENERAL INFORMATION:
; APPLICANT: Rogan, Peter
; APPLICANT: Knoll, Joan
; TITLE OF INVENTION: SUBTELOMERIC DNA PROBES AND METHOD OF PRODUCING SAME
; FILE REFERENCE: 33026-B
; CURRENT APPLICATION NUMBER: US/10/676,248B
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: 60/415,345
; PRIOR FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: 10/676,248
; PRIOR FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: 60/494,494
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 251
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 245
; LENGTH: 3026
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1843)..(1843)
; OTHER INFORMATION: n is a, c, t or g
US-10-676-248B-245
Query Match 10.2%; Score 222; DB 18; Length 3026;
Best Local Similarity 95.4%; Pred No. 1.8e-53;
Matches 250; Conservative 0; Mismatches 10; Indels 2; Gaps 2;
QY 2 TTCTGCTTCCTTTAGCGTGAA-CGCGGGTGGGTGCTCCCGTGAATAATAAATTCAC 60
Db 1758 TTCTGCTTCCTTTAGCGTGAAACGCGAGTGTGCTGCTCCCGTGAATAATAAATTCAC 1699
QY 61 CGTCACGCTTGTGTGAACGCGGGTGTTCGCGAAACTTGGAGGCTTCCCGTAAACCCG 120
Db 1698 CGTCACGCTTGTGTGAACGCGGGTGTTCGCGAAACTTGGAGGCTTCCCGTAAACCCG 1639
QY 121 CTCCTTCCTCATCTGGAGGTGGTCCCGCGGGTCCCGCTCCCTCCCTGCGCCCTC 180
Db 1638 CTCCTTCCTCATCTGGAGGTGGTCCCGCGGGTCCCGCTCCCTCCCTGCGCCCTC 1580
QY 181 CCTCTGCTGCTTTTCATTTCTCGGGCTCCGGGCGCGGAGAGCTGCATCCAGAGGA 240
Db 1579 CCTCTGCTGCTTTTCATTTCTCGGGCTCCGGGCGCGGAGAGCTGCATCCAGAGGA 1520
QY 241 GCGGTCAGAGCGGACCCGG 262
Db 1519 GCGGTCAGAGCGGACCCGG 1498
RESULT 7
US-09-997-165-7
; Sequence 7, Application US/09997165
; Patent No. US2002014199A1
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fanslow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1112
; TYPE: DNA
; ORGANISM: Mus sp.

```
Query Match      2.8%; Score 60; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 5e-07;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1668 GTGCCTCCGGGGTGGTCGGCGATCCCTCAGTCCCTCGGCCACCCGGGGGCG 1721


```
FEATURE: (1)..(5452)
LOCATION: (235)..(236)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (201)..(5252)
NAME/KEY: modified_base
LOCATION: (1)..(140)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (145)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (147)..(148)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (162)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (165)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (168)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (172)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (189)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (191)..(193)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (196)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (199)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (214)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (218)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (221)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (225)..(226)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (229)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (232)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
```

```
FEATURE:
NAME/KEY: modified_base
LOCATION: (235)..(236)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (246)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (248)..(252)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (256)..(257)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (261)..(263)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (265)..(266)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (271)..(273)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (276)..(277)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (283)..(285)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (296)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (298)..(300)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (302)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (307)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (314)..(315)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (334)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (339)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (343)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
LOCATION: (345)..(348)
NAME/KEY: modified_base
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
```

```

; NAME/KEY: modified_base
; LOCATION: (351)..(352)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (357)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (359)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (362)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (365)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (367)..(368)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (370)..(373)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (376)..(378)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base

Query Match      2.3%; Score 51.2; DB 15; Length 5452;
Best Local Similarity 20.0%; Pred. No. 0.0011;
Matches 194; Conservative 0; Mismatches 774; Indels 0; Gaps 0;

Qy 459 AGAACACCGTCATGTCCTGCAACATCTCCAAAGCCTTCTCCCATGTCAACATCAAGCTGC 518
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5270 AGGGGNCCTCAATTGTTTAANAANNCNNNTCCCTCNCCTCCCTCGNNNNNCCCTN 5211
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 519 GTGCCACGGGAGGAGCGGCATCTTCAATGAGGTGGCTCCAGGCTACTTCTCCGGG 578
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5210 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5151
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 579 ACGGCTGCAGCTCCAGGTTCAAGGAGGGTGCGCACAGCTGTGATCAAGGCGCCGG 638
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5150 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5091
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 639 ACTCCCATGCTGGGCTGTACATGTGGCACCTCGTGGGACACAGAGAAATAACAGACA 698
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5090 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5031
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 699 TCACGCTGGAGGTTTCAGGTGCAGAACCCAGTCGCGCCCTGACACTGGGTTCTGGC 758
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5030 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4971
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 759 TGCCAGCGGTGCTACTGCTCTTCACTCTTGGTCTGCTGCTGCTGCTGCTGCTGCT 818
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4970 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4911
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 819 ACAGGTGCGGCTGTTCCAGCAACCGCGGAGAGAGTCTTCTCTAGAACCCACAGA 878
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4910 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4851
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 879 TGAAGGTGCGAGCCTTCAGAGCGGAGGCCAGAGGCTGTGAGCAGAGCCTCCGCTGA 938
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4850 CCCCCCCCCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4791
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 939 TGTGACCCGACACTCCGAGGCCACCCAGGCGCTGGCACTGTGTTCAAAACCTCAC 998
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4790 CCCCCCCCCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4731
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

Qy 999 CACTTGGAGCCCTGGAGCTGTGTCCGCCCAACCCCTTTTTCATATGCGCAGACCCA 1058
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4730 CANNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4671
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1059 TAGCGGCTGCAAGGACAGAGAGGACACAGGAGAGGAGCCCTGAGTGGGAGCTTGG 1118
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4670 CCCCCCCCCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCN 4611
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1119 GCGGGGCTGGGTCTCTCGTCCACCCGAGGAGGACAGACACCGGCTGTGTTGGAGG 1178
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4610 NCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCN 4551
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1179 GGGGCTGTGTCAACCCACTCTCTGGGTGCGTGCAGACCCCTTCCCTCCACCCCGAG 1238
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4550 NCNNNNNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4491
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1239 TTCCAAGCTGTCTTCTCAGTTCCTTCCAAATGGAACACCTCAGCTCCGAGCACCAG 1298
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4490 CCCCCCCCCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCN 4431
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1299 TTACCAGGACGATGCGCCCTCTCTGCGCTCATCAAAACCCAGACAGCGGACTCCCT 1358
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4430 NNNNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCN 4371
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1359 CTGCCACCCAGGCTGTCTCGGCGCCAGGTTGTGGGTGCGGTCTCTCTCCACTCCCA 1418
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4370 NNNNNNNNNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4311
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1419 CCGGCCCC 1426
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4310 NCCNCCNCC 4303
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-10-292-798-1189/c
; Sequence 1189, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1189
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(5452)
; NAME/KEY: CDS
; LOCATION: (201)..(5252)
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (1)..(140)
; OTHER INFORMATION: a, t, c, g, unknown or other
; NAME/KEY: modified base
; LOCATION: (145)..(145)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
```

[illegible]

OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (365)..(365)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (367)..(368)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (370)..(373)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (376)..(378)
OTHER INFORMATION: a, t, c, g, unknown or other

Query Match 2.3%; Score 51.2; DB 17; Length 5452;
Best Local Similarity 20.0%; Pred. No. 0.0011;
Matches 194; Conservative 0; Mismatches 774; Indels 0; Gaps 0;
QY 459 AGAACACCGTCTGCTGCAACATCTCCAAACGCTTCTCCCATCTCAACATCAAGCTGC 518
DB 5270 AGGGGNCCTTCAATTGTTTAAANAANNCCNNTCTCNCCTCCCTCGNNNNNCCCTN 5211
QY 519 GTGCCACGGGAGGAGCGCATCTTCAATAGGTGGCTCCAGGCTACTTCTCCCGG 578
DB 5210 NNN 5151
QY 579 ACGGCTGGCAGCTCCAGGTTACGAGGAGGGTGCGACAGCTGTGTATCAAAAGCGCCGG 638
DB 5150 NNN 5091
QY 639 ACTCCCATCTGGGTGTATGTGCGACCTCGTGGGACACGAGAAATACACACAAG 698
DB 5090 NNN 5031
QY 699 TCACGCTGGAGGTTTCAGGTGCAGAACCCAGTCGCGCCCTGACACTGGGTTCTGGCCTG 758
DB 5030 NNN 4971
QY 759 TGCCAGCGGTGTCTACTGTCTTCACTCTTGGTGGCTGTGTCTGTCTGTCTGTCTGT 818
DB 4970 NNN 4911
QY 819 ACAGGTGCGGCTGTTCCAGCAACCGCGGAGAGAGTCTTCTCTCTAGAACCCGAGA 878
DB 4910 NNN 4851
QY 879 TGAAGGTGCGAGCCTTCAGAGCGGAGCCAGCAGGCTGAGCAGAGCCTCGCTGAAC 938
DB 4850 CCCCCCCCCCNNNNNCCCCCCCCCNCCNCCNCCCCCCCCCNCCNCCNCCNCCNCC 4791
QY 939 TGTGAGCCCCAGACTCCGAGCCCAACCAAGGCGCTGGCACTGTGTTCAAACCCCTCAC 998
DB 4790 CCCCCCCCCCNNNNNCCCCCCCCCNCCNCCNCCCCCCCCCNCCNCCNCCNCCNCC 4731
QY 999 CACTTGGAGCCTGGAGTGTGTGCCCCCAACCTTGTTCATATGCGCAGACCCA 1058
DB 4730 CNNNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4671
QY 1059 TAGCGGCTGCAAGGAGAGGAGACAGAGAGGAGCAGCCTGTAGTGGGCTG 1118
DB 4670 CCCCCCCCCCNNNNNCCCCCCCCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4611
QY 1119 GCGGGGCTGGGTCTCTGTCCTCAACCGAGGGGACAGACACCGGCTTGTGGAGGCT 1178
DB 4610 NCCCCCCCCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4551
QY 1179 GGGCTCTGTGTACCACTCTGTGGGTGGTGACACCTTCTCCCTCAACCCCCAGGTC 1238
DB 4550 NNNNNNNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4491

QY 1239 TTCCAAGCTCTGCTTCTCAGTTTCCAAAATGGAACCACTTCACTTCCGAGCACCAGAC 1298
DB 4490 CCCCCCCCCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4431
QY 1299 TTACAGGACGATGCGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTT 1358
DB 4430 NNNNNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4371
QY 1359 CTGCCACCCAGGCTGTGTCGCGCCCGCAGGTGTGGGTCTGCTTCTTCACTTCCAGGGCT 1418
DB 4370 NNNNNNNNNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 4311
QY 1419 CCGCGCCC 1426
DB 4310 NCCCCCCC 4303
RESULT 14
US-10-424-599-61903
; Sequence 61903, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 61903
; LENGTH: 805
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(805)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_26913C.1
US-10-424-599-61903

Query Match 2.3%; Score 50.4; DB 17; Length 805;
Best Local Similarity 47.2%; Pred. No. 0.00086;
Matches 141; Conservative 0; Mismatches 158; Indels 0; Gaps 0;
QY 1420 CGCGCCCAAGTGAAGGGGCCCCCTGCGGAGGCTTCAGACACACTGAGGTTTCAAGGGCTGGGG 1479
DB 411 CGGGCGGGGTCCCGCGCNCNCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 470
QY 1480 GGGCTTGGACATACCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1539
DB 471 CGGGCGGG 530
QY 1540 GCGCGGG 1599
DB 531 GGG 590
QY 1600 GCCAGGGGGGCAAGCGCAGCGCGGGCTGGGAGCGCGCGCGCGCGCGCGCGCGCGCG 1659
DB 591 GGG 650
QY 1660 AGGTGGAGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1718
DB 651 NGGG 709

RESULT 15
US-10-184-644-346/c
; Sequence 346, Application US/10184644

Sun Mar 20 14:17:44 2005

```
Publication NO. US20030044930A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184,644
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 346
LENGTH: 671
TYPE: PRT
ORGANISM: Homo Sapien
US-10-184-644-346

Query Match      2.3%; Score 50.2; DB 14; Length 671;
Best Local Similarity 6.8%; Pred. No. 0.00091;
Matches 39; Conservative 208; Mismatches 329; Indels 0; Gaps 0;

Qy 927 CTTCCGCTGAACCTGGACCCAGCTCCGAGCCACCCAGCGCGCTGGCACTGGTGT 986
Db 648 H.H.C.YBHD.H.SSGY..D.MH.TS.CM.SAT.CNB.TS.HAS.MH.TM 589
Qy 987 TCAACCTCACCACCTGGAGCTGTGCTCCGCCCAACCTGTTCATAT 1046
Db 588 ...T..CT...CT.M.TM.M.KTNBM.BT..M.C.BYSM.YARB.TTM.KBHMNTMR 529
Qy 1047 GCCGACAGCCATGACCGCTGTGACGAGGAGGACAGGAGGAGCCCTGAGTGC 1106
Db 528 YY.MM.ATBBHANM....SAB..AC....TN.GYMBH.SHB.M.TM..SD.M..M..BS. 469
Qy 1107 CGACCTTGGTGGCGGCTGGTCTCTGCTCCCGAGGCGGACACACCGCTT 1166
Db 468 .MM.BMYHS.C.SYMYAY.BMB.MT..M..YBYM.M.C.MYT..ASSS..MNTSMWT 409
Qy 1167 GCTTGGCAGGCTGGGCTCTGTGTACCCACTCTGGGTGGTGCGACACCTTCCCTCC 1226
Db 408 Y...T...SHSS..CYCYMYCYMBB..HH.Y..HCSSCCCY.T.CYHTY.YYY.M. 349
Qy 1227 ACCCCCGAGGCTTCCAGCTCTGCTTCTCAGTTTCCAAATGGAACACCTCACCTCC 1286
Db 348 .M...Y.YY.....YYTYT.HYTMW.S.YHB.HSHSHSSSS..Y..M.MYC 289
Qy 1287 GCAGCACCAGCTTACACGAGGATGCCCTCCCTCTGCTCCCTCATCAACCCACAGACC 1346
Db 288 Y.M...M.T.MYCY..MMBSHSHSSSSSSSSSTSTYKTB...MTCSHSTMSSHSTS. 229
Qy 1347 CGGACTCCCTTCTCCACCCAGGCTGTGTCGGCCCGAGGTGGGTGGCTCTCTCC 1406
Db 228 TMMMYCC..CYI.TYBTMM..A.H.HSAM.S.SSS..SN...S.SBST.H.HSSTMYTY 169
Qy 1407 ACTCCCGGCTCCGCGGAGTGGGCGGCTGCGGAGGCTCAGACACACTGGAG 1466
Db 168 MSBSKM.T.AMTM.CSNHSSMSSHSHS.KYHSTTATATATKBCYH.HH.HTHSCTH.N 109
Qy 1467 TTCAGGCTGGGGGCGCTTGGCATACTCTCTCC 1502
Db 108 T..T..SHSSSB...TSRSTD.NN...W.C.N.CMYM 73
```

Search completed: March 20, 2005, 09:46:02
Job time : 1262 secs